IRISPen II

(EXECUTIVE)

U S E R ' S
G U I D E

I.R.I.S.

Document to Knowledge™

INTRODUCING THE **IRISP**EN

The IRISPen may be the biggest sensation in computerland since the mouse! As efficient as computers are, you have to encode your data first. If you've ever rekeyed text, you know how boring and time-consuming it can be....

The IRISPen, a unique product for every computer user, offers instantaneous text retrieval at your fingertips. Text recognition where you want it, when you want it.

The IRISPen consists of two components: a unique greyscale pen-sized OCR scanner and the powerful PCR-II ("Pen Character Recognition") text recognition technology.

The basic principle could hardly be easier: slide the IRISPen over a line of text and it is entered into your application instantly, as if keyed in but up to 30 times faster. Up to a 100 characters can be entered per second. Cut from paper, paste to Windows©: can you imagine a simpler way of entering text?

Imagine you have to write a report. The IRISPen copies excerpts from books, magazines, newspapers and typed or printed matter, and enters them directly in your wordprocessor at the cursor position.

Encode last minute data from the stock exchange into your notebook before you have that brainstorm with your director. And why not enter your appointments in your office software automatically? Rule the IRISPen over your visitor's business card to put him in your personal organizer. Or fix that meeting right away in your electronic agenda.

Encoding checks, invoices, bank statements etc. becomes a cinch: the IRISPen copies your numeric data directly into your database, spreadsheet and bookkeeping software. Recognition can be limited to numeric characters and customized character sets.

The IRISPen offers many other advanced features for professional applications: character conversion, automatic execution of keyboard commands, detection of table cells etc. The IRISPen is undoubtedly the most advanced data entry solution available!

Since the IRISPen connects to the computer's USB port without the need for an external power supply, you have a portable OCR system, an ideal solution, even for notebook owners.

Finally, the IRISPen allows you to scan small graphics: logos, signatures, handwritten notes etc. are inserted directly in your documents by sliding the pen scanner over them.

ON TO THE IRISPEN EXECUTIVE!

The IRISPen Executive software converts your IRISPen into a multifunctional productivity tool. This software-only upgrade extends the capabilities of the pen-sized OCR scanner dramatically; as a result, your IRISPen becomes unquestionably the most versatile reading device available!

Up to now, you were able to recognize printed and typed material. The Executive software extends your system with four major add-on capabilities: speech synthesis, bar code reading, handprinting recognition and recognition of dot matrix documents. Furthermore, the IRISPen Executive is equipped with advanced image adjustment features that allow to read a wider range of documents and a multiline reading mode that increases the pen scanner's ease of use.

The **speech synthesis** capability "Read & Speak" opens the door for multimedia applications and allows "hands- and eyes-busy" data entry as your sound card reads out all recognized text. You no longer have to look up to the computer screen to verify the results: the IRISPen results are rendered by a human voice, the most natural way of interfacing with a machine!

A single speech library is supplied standard; text-to-speech modules for additional languages can further extend your system.

Bar code reading opens a world of new applications to IRISPen owners as their pen scanner becomes an ergonomic, light-weight and compact bar code scanner for the office environment. Information desks, marketing surveys, cus-

tomer identification, order picking and many other examples of high accuracy data entry can benefit greatly from this bar code reading solution.

Although all widespread bar code standards or "symbologies" are supported, you don't have to specify the bar code type: the user-friendliness is such that the IRISPen automatically detects which bar code standard is used! The IRISPen Executive offers transparant bar code reading into any Windows application. No programming or parameter setting is involved in its operation, its use is automatic and immediate.

You can even use a **mixed reading mode** where bar codes are read alongside other data - for instance "normal" text. The IRISPen "auto-detects" which type of data is being read. For instance, you can scan address labels where the printed data is complemented by a bar code without ever changing the reading mode.

Encoding forms and labels can also benefit greatly from numeric **handprinting recognition**. This proprietary recognition technology supports all natural writing styles - American and European; contrary to many other handwriting recognition solutions, no training phase or imposed style is required.

And there's another reading capability added: **recognition of dot matrix** printed documents. "Draft" or "9 pin" dot matrix symbols are made up of isolated, separate dots, and require highly specialized recognition routines to recognize them.

Advanced **image adjustment** features optimize the images for the OCR process: as the scanning brightness gets adjusted to the color tones of your documents, documents that are otherwise illegible can be recognized. "Despeckling" removes "noise" from the scanned images, again resulting in higher OCR accuracy.

Finally, the **multiline scanning** mode increases the ease-of-use. It allows you to scan several lines of text. All scanned data is temporarily "buffered" until you press the pen button to "trigger" the output. This scanning mode also ensures a natural prosody. When the speech synthesis has to pronounce individual text

lines, rather than full sentences, the intonation may be somewhat less than "human".

How Does the IRISPen Work?

The pen-sized scanner functions as a greyscale camera that sends an image to the computer. To recapture printed or typed information, you simply rule the pen over a line of text. You start scanning as soon as you press the IRISPen onto the paper; the scanned text is lit up by the scanner. As you scan the text, an image consisting of intense points (so-called "pixels") on a lighter background is generated and transmitted to the PC.

When you stop scanning by lifting the pen from the paper, the software takes over. Intelligent routines convert the greyscale images into black-and-white images. The text recognition software converts these images into editable characters and sends the recognized text promptly to the target application where it is inserted at the current cursor position.

As you are operating a "hand-held" scanner, you will inevitably get wavy, distorted images. The highly specialized character recognition software PCR-II is nevertheless able to cope beautifully with this low-quality "material".

This proprietary technology was designed specifically for pen scanners. It allows real-time text recognition despite wavy, low-contrast and poor quality images containing highly distorted font shapes, and incorporates advanced image processing techniques and unique artificial intelligence routines to this end.

Table of Contents

USER'S GUIDE

Chapter 1: Installation

Chapter 2: Guided Tour

Chapter 3: Read & Speak

Chapter 4: Bar Code Reading

Chapter 5: Handprinting Recognition

Chapter 6: Command Reference

*U*SER'S *G*UIDE

Appendix B: Syntax Rules Speech Synthesis

Appendix C: Bar Code Symbologies

Appendix D: Handwriting and Form Design Rules

CREDITS AND COPYRIGHTS

The IRISPen, the IRISPen Executive and the IRISPen Translator software were designed and developed by I.R.I.S. OCR and Linguistic technology by I.R.I.S. I.R.I.S. detains the copyrights to the IRISPen software, the IRISPen Executive software, the IRISPen Translator software, Read & Speak, the PBR ("Pen Bar Code Reading") technology, the PCR-II ("Pen Character Recognition") technology, the PHR ("Pen Handprinting Recognition") technology, the linguistic technology, the on-line help system and this manual.

The I.R.I.S. logo, the IRISPen, PBR, PCR-II, PHR, I.R.I.S. Linguistic Technology and Read & Speak are trademarks of I.R.I.S.

The L & H logo and L & H Quality Speech are trademarks of Lernout & Hauspie.

Excel, Outlook, Windows and Word are registered trademarks of Microsoft Corp. Intel is a registered trademark of Intel Corp.

Chapter 1

INSTALLATION

This chapter discusses the system requirements and installation of the IRISPen, the IRISPen Executive and the IRISPen Translator. The IRISPen consists of a hardware component and a software package.

SYSTEM REQUIREMENTS

This is the minimal system configuration required to use the IRISPen:

❑ a 486 based Intel PC or compatible. A Pentium based PC is recommended.

❑ 16 MB RAM.

❑ 50 MB hard disk space for the Standard software, 60 MB for the Executive software and 80 MB for the Translator software. Substract 10 MB when you leave the electronic manual on the CD-ROM.

❑ a 256 color monitor. It is recommended that you dispose of a 256 color monitor to display the greyscale images generated by the pen scanner.

❑ the Windows ME, 2000 or 98 operating system.

❑ a free USB port to connect the IRISPen scanner.

The **speech synthesis** of the IRISPen Executive and the IRISPen Translator requires that your PC is equipped with a Windows compatible **sound card**. For quality sound, preference is given to a 64-bit sound card. All Windows compatible audio cards are supported - Creative Labs SoundBlaster and compatibles etc.

To install additional **speech modules**, 5 MB of free hard disk space is required per language. Ten languages are currently available: English and British

English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.

To install additional **translation pairs**, 15 MB of free hard disk space is required.

Contact your computer reseller if necessary.

WHAT YOU SEE IS WHAT YOU GET

The IRISPen product package should contain the pen scanner, the IRISPen CD-ROM, this manual, a registration card and a collection of IRISPen test sheets.

Contact your I.R.I.S. reseller if any items are missing from your package.

INSTALLING THE IRISPEN SOFTWARE

The IRISPen software is delivered exclusively on an **autorunning CD-ROM**. To install, simply insert the CD-ROM in your CD-ROM drive and wait for the installation program to start running. Follow the on-screen instructions.

Should the installation not begin to run when the CD-ROM is inserted in your CD-ROM drive, run the setup program MENU.EXE to install the software.

Users of Windows 2000 must ensure that they have the necessary **access rights** - contact the system administrator if necessary.

Some installation options are offered. Be sure to install the Acrobat Reader software required to access the software documentation, should this be necessary. The **electronic manual** is by default copied to your hard disk. You can also leave it on the CD-ROM.

Install the **linguistic databases** of all languages you intend to read. By default, all lexicons are installed.
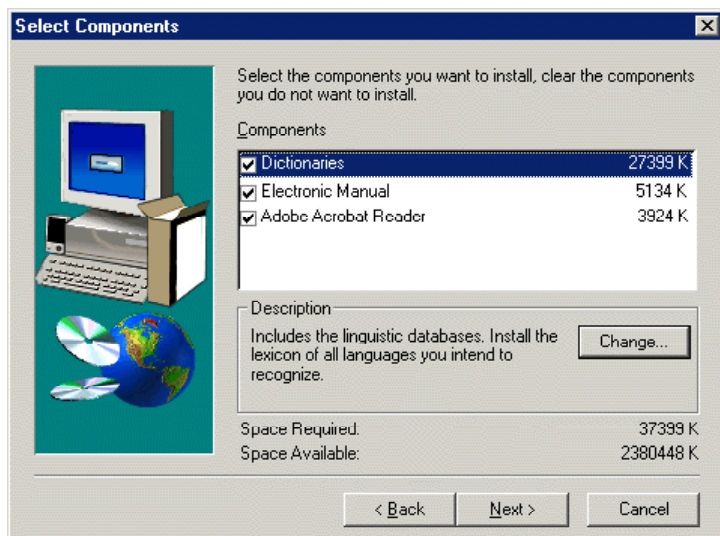
For the IRISPen Translator, you also have to select a language pair for the bidirectional **translation**. For instance, the IRISPen Translator English-French translates English into French and French into English. IRISPen Translator owners are entitled to one language pair. Extra language pairs can be purchased separately from I.R.I.S.

Select Sub-components

Select the components you want to install, clear the components you do not want to install.

Sub-components

☑ English-French	11875 K
☑ English-German	17365 K
☑ English-Italian	12571 K
☑ English-Portuguese	11636 K
☑ English-Spanish	10997 K

Description

Is required to perform translation from English to French and from French to English.

Space Required: 118064 K
Space Available: 2354384 K

[Continue]

IRISPen Executive owners are entitled to one speech library of their choice; hence, they will be invited to select the language of the **speech synthesis**. The IRISPen Translator is equipped with the speech libraries of both languages in the translation pair. Extra speech modules can be purchased separately from I.R.I.S.
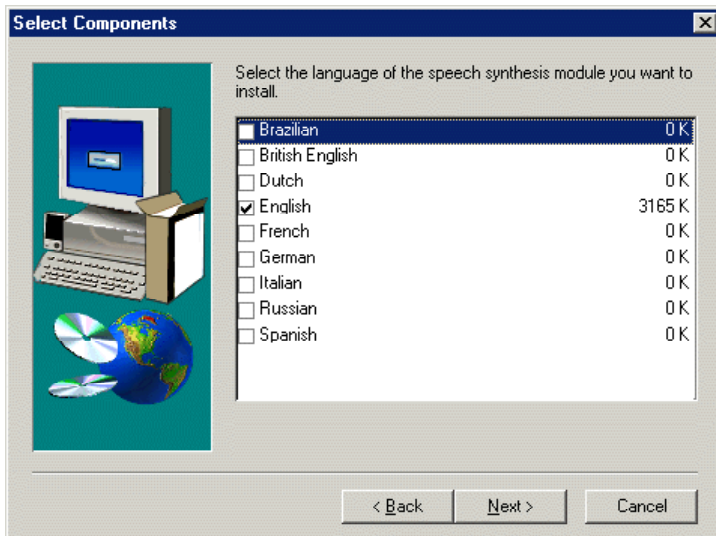
The submenu "I.R.I.S. Applications - IRISPen" under the "Programs" menu is created automatically by the installation program.
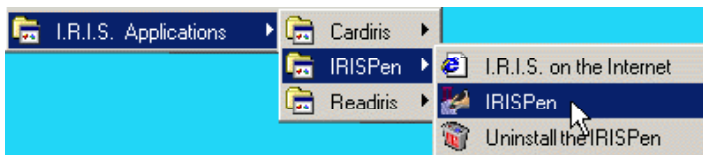
The installation program also creates a **shortcut** for the IRISPen application on the Windows desktop. As a result, you are able to start the IRISPen software directly from your desktop.
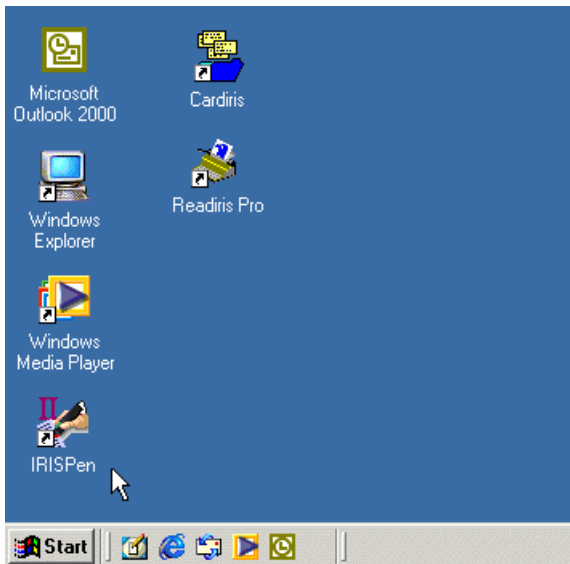
INSTALLING THE IRISPEN HARDWARE

The IRISPen uses a **USB connection**. Standard PCs are equipped with two USB ports; this capability can be extended by installing a USB "hub" on your PC. Note above all that USB support requires the operating system Windows ME, 2000 or Windows 98!

Secondly, you are recommended to install the software *before* the hardware: as soon as you connect a USB device to your PC, the operating system starts looking for appropriate **drivers** - which are installed automatically by the installation program.

The **hardware installation** of the pen scanner couldn't be easier: you can "hot plug" the pen scanner into the PC's USB port without having to reconfigure or reboot your PC.
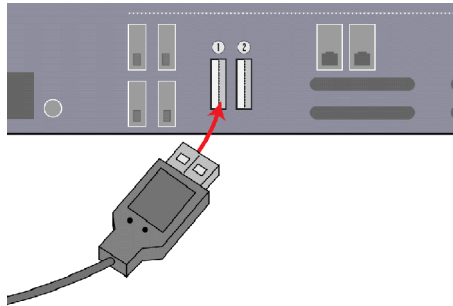
The use of the USB port also implies that no external power supply is required for the pen scanner: the PC supplies the necessary power via the USB port!

UNINSTALLING THE IRISPEN SOFTWARE

There are only two correct ways of uninstalling the IRISPen: using the IRISPen "uninstall" program and using the Windows (un)install wizard. You are strongly recommended *not* to uninstall the IRISPen software or its software modules by manually erasing the program files.

IRISPen "uninstall" program

Select "Uninstall IRISPen" under the submenu "I.R.I.S. Applications - IRISPen" to start the IRISPen "uninstall" program and follow the on-screen instructions.
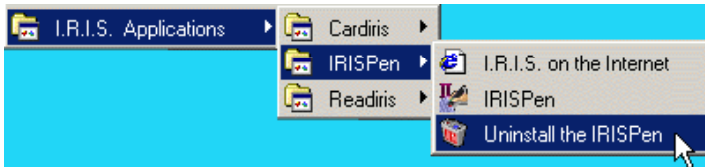
Windows (un)install wizard

Execute the following steps to make use of the Windows (un)install wizard.

❑ Click "Settings" under the "Start" menu of Windows and go to the "Control Panel".
❑ Click the icon "Add/Remove Programs" under the control panel.
❑ The IRISPen is listed under the "Install/Uninstall" tab.

USER'S GUIDE
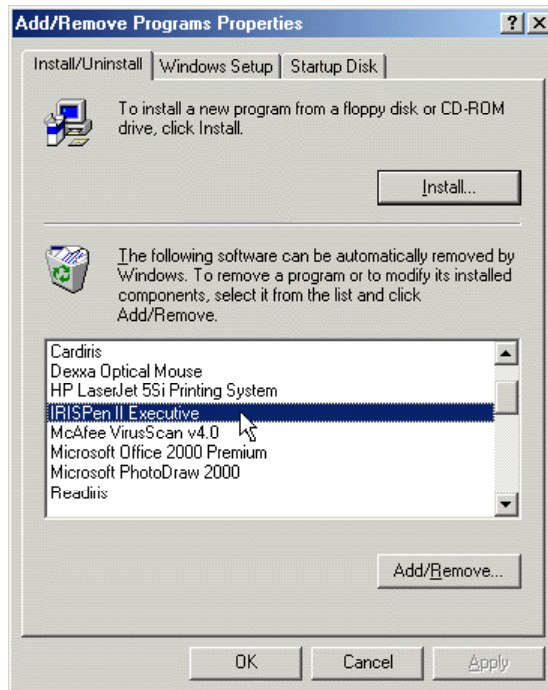
❑ Select IRISPen and click the "Add/Remove" button to remove the IRISPen software.
❑ Follow the on-screen instructions.

REGISTER TO VOTE!

Don't forget to register your IRISPen licence! Doing so will allow us to keep you informed of future product developments and related I.R.I.S. products. The

registration benefits, including free product support and special offers, are strictly limited to registered users.

You can register in many ways: by sending in your registration card or faxing its electronic counterpart, by calling I.R.I.S. during working hours and by filling out a registration form on the I.R.I.S. web site!
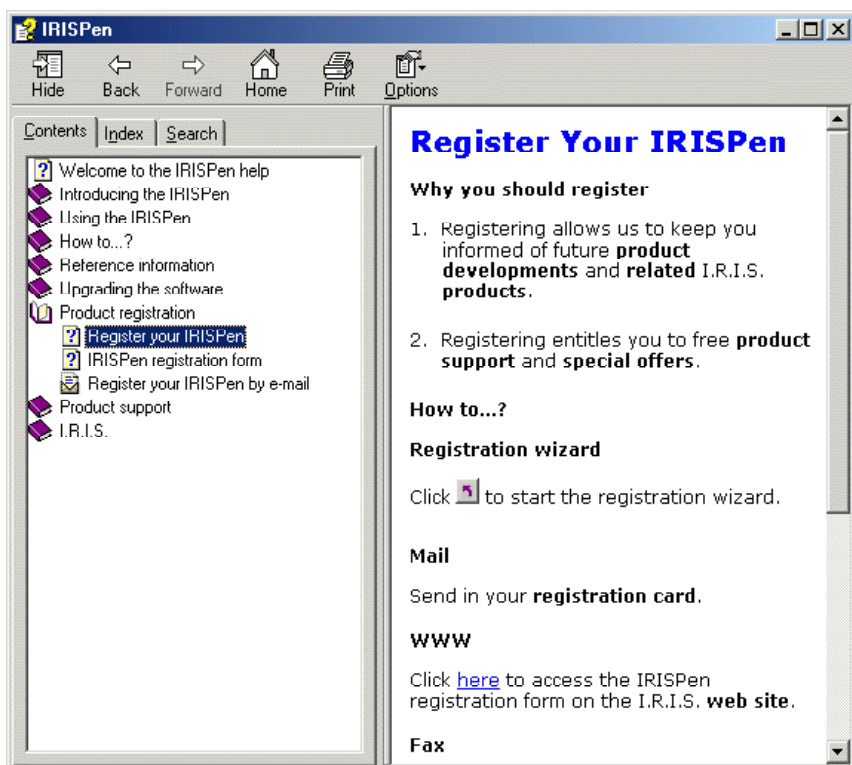
The IRISPen **registration wizard** as you'll find under the menu "Register" of the IRISPen software can guide you through the registration process comfortably.

GETTING PRODUCT SUPPORT

The command "Product Support" under the "Help" menu of the IRISPen software details how you can get technical support. Please describe the phenomenon you experience clearly and include all relevant data concerning your IRISPen and your computer system.
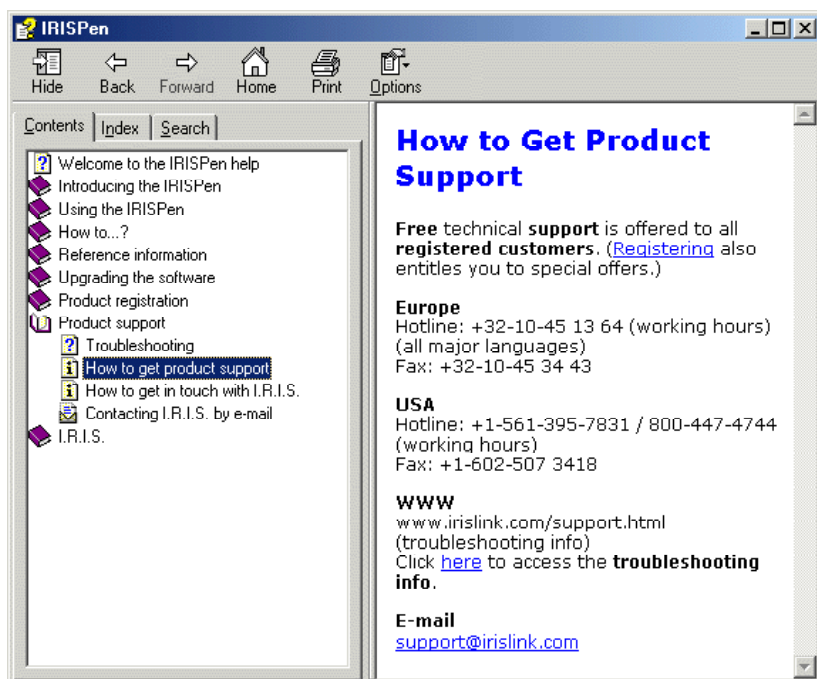
GETTING IN TOUCH WITH I.R.S.

You can contact I.R.I.S. to learn more about complementary software and related products; the command "Contact I.R.I.S." under the "Help" menu of IRISPen details in which ways you can get in touch with I.R.I.S.

USER'S GUIDE
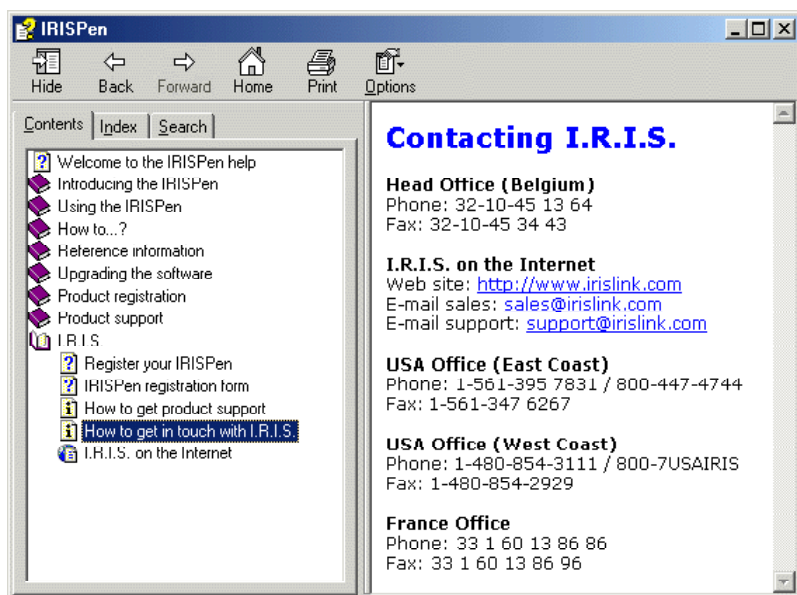
An application icon in the submenu "I.R.I.S. Applications - IRISPen" under the "Programs" menu takes you directly to the I.R.I.S. **home page**. So does the command "I.R.I.S. on the Internet" under the "Help" menu of the IRISPen software.
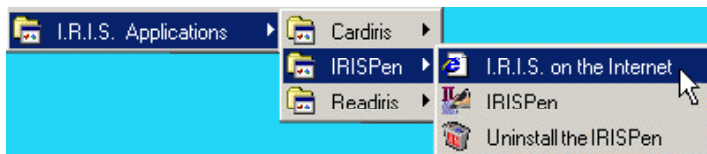
Chapter 2
GUIDED TOUR

The IRISPen is a sophisticated product equipped with many advanced features for professional applications. Although it is impossible to explore every imaginable application, we will discuss in this chapter the major applications and add many tips and hints concerning the use of the IRISPen.

Let's have a look at the pen scanner first. What should you know about the hardware and how do you use it?

HAVING A LOOK AT THE PEN SCANNER

The IRISPen is equipped with a roller which functions as **start switch**: when you press the IRISPen onto a piece of paper, you turn it on, when you lift the pen scanner away, the scanning ends. (The start switch only comes on if you hold the IRISPen at the correct angle.)

As you scan data line by line, you should always be able to see what you are scanning. The **view window** allows this. As soon as the start switch is activated, a red beam from the view window lights up the area you're scanning.

The scanning can also be accompanied by **sound**: an option activates your computer's speaker when you start and stop scanning and when you send the recognized text to the output target. So, you will hear three system beeps for each scan.

Whether you do so, depends on the "Sound" button on the toolbar (or the option "Sound" under the "Options" menu).

This option is not compatible with the speech synthesis of the IRISPen Executive and the IRISPen Translator: you cannot produce system beeps and pronounce the recognized text at the same time!

View Window   Line Guide

Start Switch

Pen Button
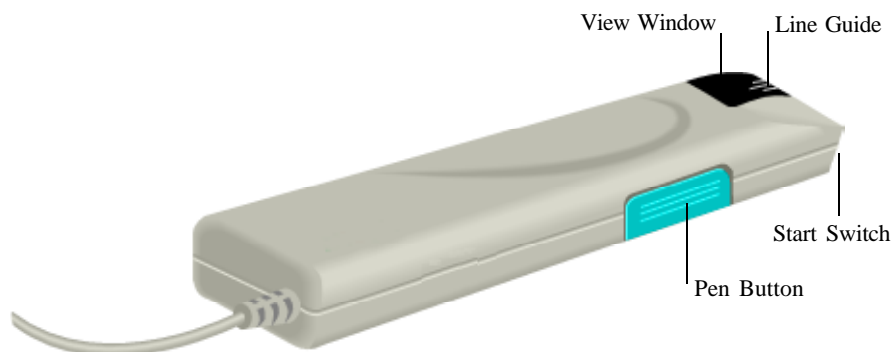
The **line guide** on the view window helps you scan accurately. Steady scanning not only improves the recognition but is specifically useful when you want to read text with limited interline spacing. The line guide allows you to isolate the correct information: the IRISPen recognizes the text line closest to the line guide.

The **pen button** is a programmable command button: when you press the button, you can execute a keyboard command, for instance <Enter>, add characters such as "$" etc. As you are scanning, you can execute necessary keyboard commands directly with your IRISPen. As is the case with a computer mouse, the single and double-click of the pen button correspond to different commands.

USER'S GUIDE

STARTING THE SOFTWARE UP

So much for the pen scanner. Let's get started with the software: now that we have installed the IRISPen system, we want to start scanning text into our applications.

To start the IRISPen software, click on the IRISPen application in the submenu "I.R.I.S. Applications - IRISPen", or use the shortcut created by the installation program on the Windows desktop.

The IRISPen startup screen appears for a few seconds. It displays the version and copyrights of the IRISPen software. It also gives direct access to the I.R.I.S. **home page** - simply click on the URL www.irislink.com to visit the I.R.I.S. web site.
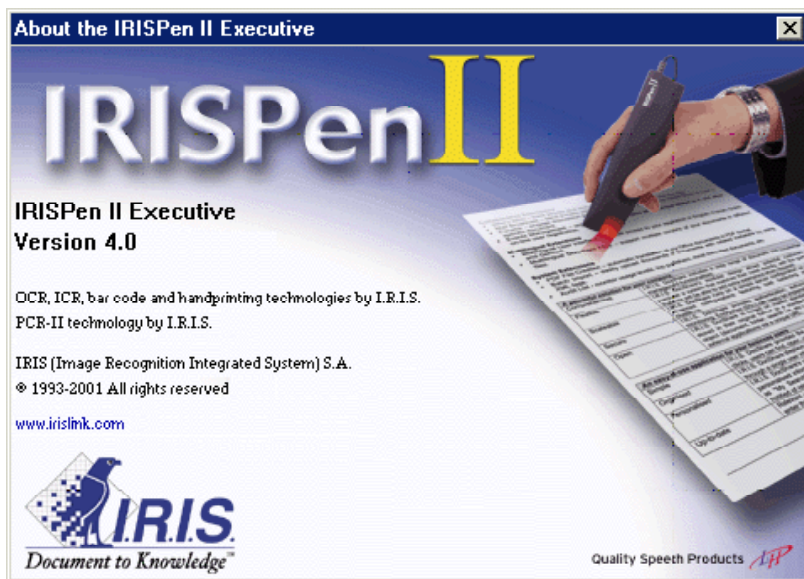
Next, the IRISPen application window and two additional windows, the (greyscale) image window and the status window, are displayed.

We'll concentrate on the IRISPen application window first. It consists of three parts: the *menu bar* comprises all software commands, the *toolbar* gives quick

access to all frequent commands and the *window pane* contains the basic settings.

SETTING UP THE NECESSARY PARAMETERS

Let's have a look at the basic IRISPen settings in the application's window pane.

Reading Mode

Determines the reading mode: alphanumeric, bar codes, black-and-white bitmap to clipboard, greyscale bitmap to clipboard, handprinting and numeric. The options "Bar Codes" and "Handprinting" are limited to the IRISPen Executive.
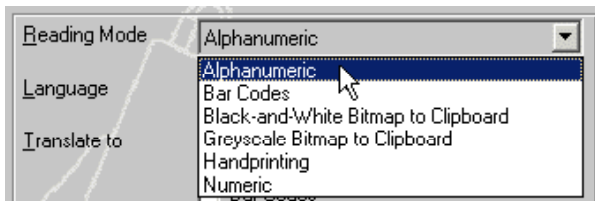
Select "Alphanumeric" for the time being.

Language

Indicates the language of the text you want to recognize. The IRISPen supports no less than 55 languages!

Select "English" for now.

Translate to (IRISPen Translator)

Indicates the target language of the translation: you scan for instance English text and translate it into French. Sentence translation is obviously limited to the IRISPen Translator.
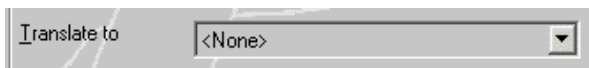
Translate to	<None> ▼

Bar Codes (IRISPen Executive)

Enables a mixed reading mode where bar codes are read alongside other data. The parameter "Reading Mode" sets the primary reading mode, this option determines whether you read bar codes *at the same time*.

☑ Bar Codes

Simultaneous bar code reading is limited to the IRISPen Executive.

Leave this option disabled for the time being.

Dot Matrix (IRISPen Executive)

This option determines whether you are recognizing "normal" or dot matrix printed documents. Dot matrix recognition is limited to the IRISPen Executive.

☑ Dot Matrix

Leave this option disabled for the time being.

Output Target

The recognized text can be inserted directly in the active application or transferred to the clipboard for later pasting.

Output Target	○ Clipboard	● Application
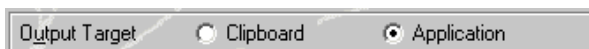
Obviously, the clipboard can only contain the recognized text of the last scan!

A Windows application is mostly on the receiving end. Select "Application" here.

Multilines (IRISPen Executive and Translator)

The option "Multilines" sees to it that multiple lines are "buffered" before they get inserted in the target application. You can for instance send an entire paragraph to your Windows application at once. It takes a click on the pen button to "flush" the scanned text.

☑ Multilines
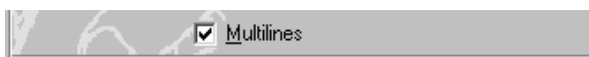
This option is limited to the IRISPen Executive and the IRISPen Translator.

ABOUT THE STATUS WINDOW...

The basic settings are displayed in the IRISPen **status window**. Minimally, this window lists the reading mode, the selected language and the output target.

Alphanumeric, English, Application

The status window also indicates more advanced settings: whether the dot matrix reading is enabled and whether the simultaneous bar code reading is active.

Alphanumeric-Bar Codes, English, Dot Matrix, Application
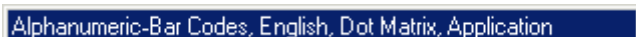
For the IRISPen Translator, the translation direction is indicated. You translate for instance English text into French.
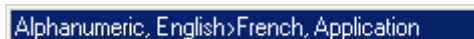
Alphanumeric, English>French, Application

Whether the status window is displayed or not depends on the option "Status Window" under the "Options" menu. It is enabled by default. Don't hesitate to drag the status window to another position on the screen!

✔ Status Window

Click the "OK" button to minimize the IRISPen application; clicking the "Minimize" button in the upper right corner amounts to the same thing. The IRISPen application now runs in the background.

... AND THE IRISPEN CONTROL ON THE TASKBAR

The Windows taskbar confirms that the IRISPen application is running: you'll find the IRISPen control on the Windows taskbar!

The IRISPen control indicates whether the pen scanner is available - if it is detached from your USB port, the Windows taskbar will indicate this.

You can call up some frequent commands by *right*-clicking the IRISPen control: amongst others, you can restore the IRISPen application, enable the status window and the greyscale image viewer. More about the greyscale image viewer shortly.
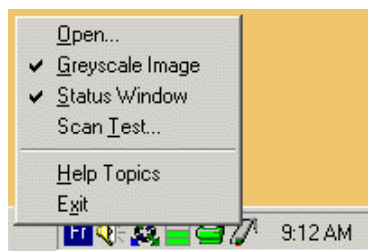
PUTTING PEN TO PAPER

Start WordPad or another text-based application. Keep your I.R.I.S. test sheet to hand.

Hold the IRISPen like a pencil. When you press the pen scanner onto the paper, you activate the start switch and the data you want to scan is illuminated through the view window by a red light. The design of the IRISPen is such that the start switch only comes on when you hold this device at the correct angle.
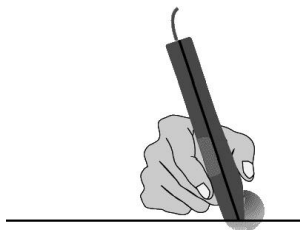
Rule the IRISPen over a text line of the I.R.I.S. test sheet at a steady rate. As soon as you lift the IRISPen, the recognized, editable text appears in your text editor or text application.

Left-handed users can adapt the scanning direction for optimal comfort. By default, the pen scanner is operated from left to right, but when you enable the option "Left-Handed Scanning" under the "Options" menu, you can slide the IRISPen in a right-to-left direction.

This only works when you hold the pen scanner rotated by 180° so that the pen button faces towards you, while the document itself retains its normal orientation! The software "compensates" the scanning direction by rotating the image by 180°.

(Alternatively, when you rotate the page *and* the pen scanner by 180°, you can disable the option "Left-Handed Scanning" and still scan right-to-left.)

Try to maintain a straight line and constant speed during your scan. The more regularly you scan, the better the recognition result will be.

You can scan some 3.15" (8 cm.) per second. When you scan too fast, image information gets lost, and the status window blinks for a few seconds inviting you to scan slower.

Maximal scanning speed exceeded: please scan slower

The IRISPen reads the text line directly under or closest to the **line guide**. Lift the IRISPen to stop scanning. As soon as you do, the recognition software analyzes the image and the scanned text appears instantly in the output target. You now have editable characters in your text editor.

ACQUIRING SLEIGHT OF HAND

During the first scans, you may not be satisfied with your results. This is undoubtedly because you do not scan comfortably. For some users, the IRISPen takes a little getting used to.

The greyscale image window and the scan test mode allow you to see the image you generate by scanning. Use these viewers to judge the quality of the images you produce and improve your scanning skills.

Let's start with the **greyscale image** window. This small window displays the "raw" greyscale images as they are generated by the pen scanner. Let's not forget that the IRISPen is the first greyscale pen scanner!

IRISPen - Greyscale Image

With quality also comes choice. We have a huge range of finishes and styles -
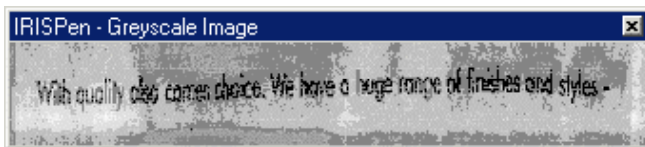
This window is enabled by default; you can activate it by right-clicking the IRISPen control on the Windows taskbar and selecting the command "Greyscale Image", and you can select the command "Greyscale Image" under the "Options" menu of the IRISPen software. Click the close button to close it.
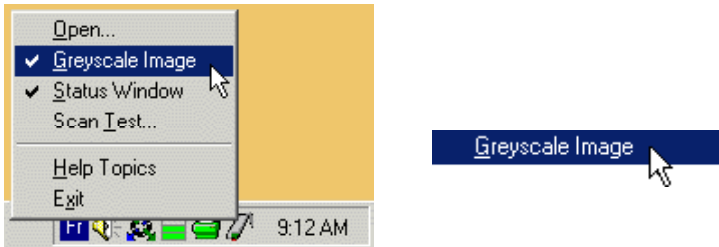
As is the case for the status window, you shouldn't hesitate to drag the status window to another position on the screen. (The position of the status window and greyscale image window are maintained when you save the default settings.)

To access the scan test window, right-click the IRISPen control on the Windows taskbar and select the command "Scan Test". Or select the "Scan Test" button on the IRISPen toolbar or even the command "Scan Test" from the "Options" menu of the IRISPen software.
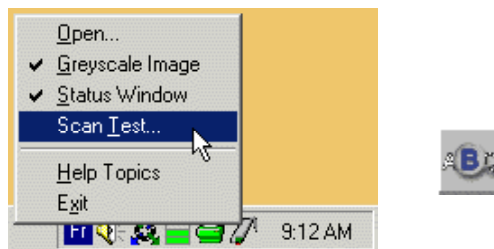
The image you generate will now be displayed in conjunction with the recognized data. The **scan test** window displays a black-and-white version of the scanned image, while the image window displays the actual - greyscale - scan.

The option "Fit to Window" ensures that the entire image is displayed in the scan test window, whatever the length of the scanned line may be.

This option is enabled by default. With this option disabled, the image isn't compressed horizontally when the scanning is completed. Use the scroll bar to display various parts of the image successively. The example below offers an illustration: a long line was scanned as the recognized text indicates, nevertheless only a small portion of the image is displayed.

Scan Test

This sample documer

This sample document shows how the IRISPen accurately recognizes a wide variety of

Brightness
- ⦿ Automatic
- ○ Manual   ☑ Smoothen Image

lighten   127   darken

Despeckle: off

0   20

☐ Fit to Window   Load Image...
☑ Recognize   Save Image...

OK   Cancel
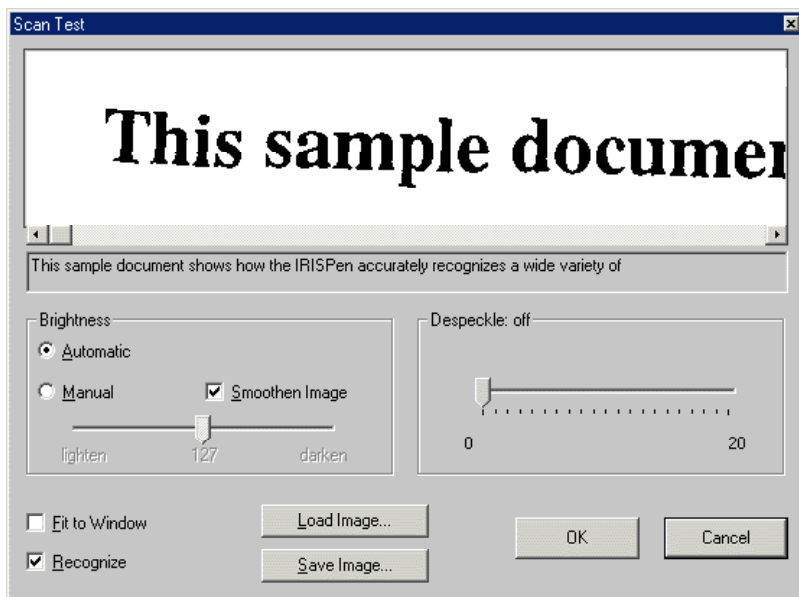
The line length is limited to 12" (30.5 cm.) per scan. The pen scanner simply shuts down when you scan a longer line.

When the scanned line is very long, you may have to scroll to display the recognized text entirely. Place the cursor in the text zone and press the End key to go directly to the end of the recognized text; press Home to return to the beginning.
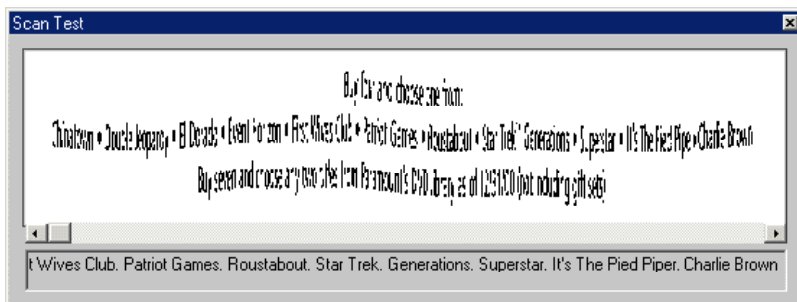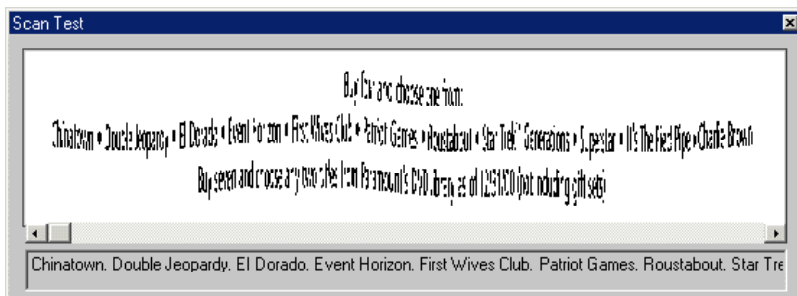
To obtain an optimal image, slide the pen smoothly and evenly over the paper. Do not lift the IRISPen away from the paper, as this ends your scan. These are not unrealistic conditions. Acquire some experience with your IRISPen and all will come naturally; you can use the I.R.I.S. test sheets to get some training.

Once you have developed the necessary "sleight of hand", the agility of the IRISPen is such that you can scan text from books, bottles etc., very much like you use a bar code reader. This kind of flexibility cannot be obtained with a sheetfed or flatbed scanner!

As soon as you have to scan text from physical, non flat objects, the IRISPen offers a unique solution. The IRISPen as a "3D" OCR product, you might say... Or you could call the IRISPen the first true handy scanner!

THE IRISPEN AND DIRECT IMAGE SCANNING

Clicking the button "Save Image" saves the scanned image in a graphic file. Three file formats are available: black-and-white and greyscale Windows bitmaps (BMP) and a proprietary black-and-white graphic format called "IRISPen image". IRISPen graphics should be used for diagnostic purposes only.
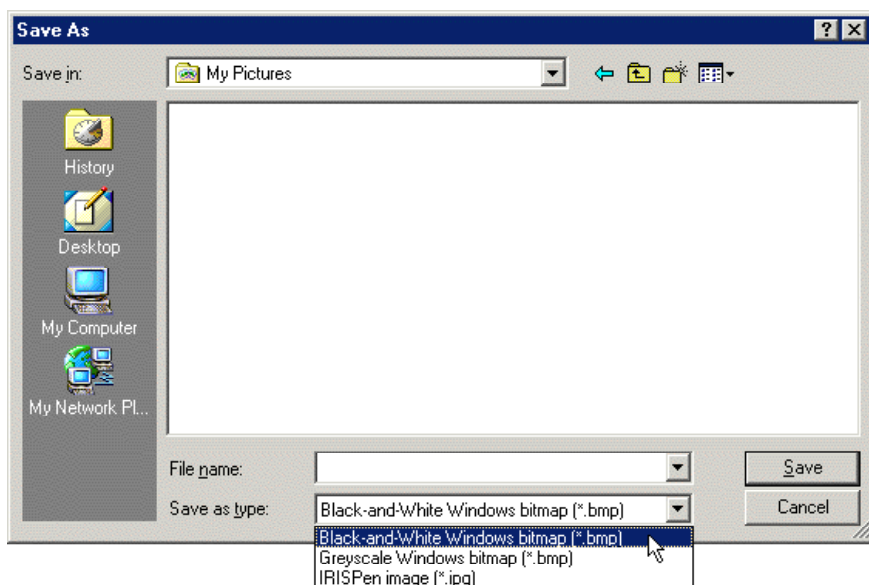
Windows bitmaps can also be used to save images: use them to scan small **graphics** such as logos, signatures etc.
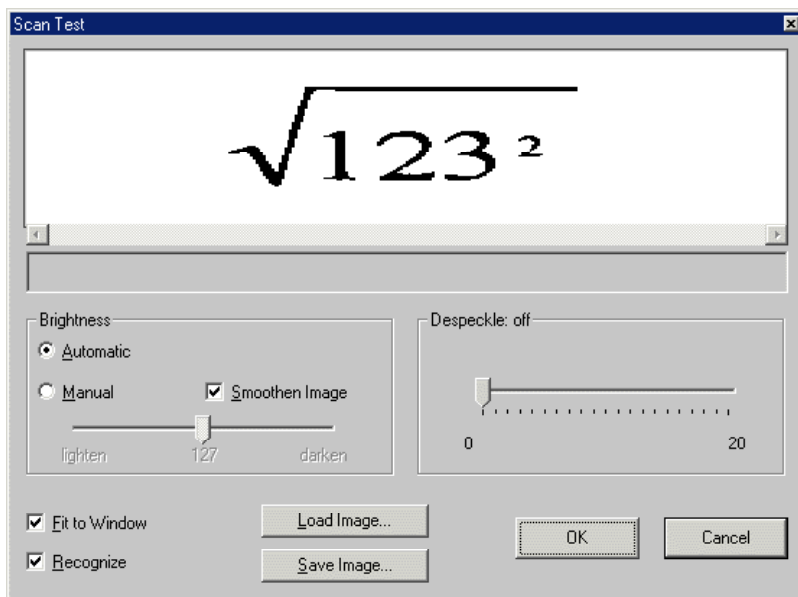
Bitmaps and IRISPen graphic files can be loaded into memory with the button "Load Image".
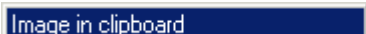
The option "Recognize" sees to it that all images you scan or open are recognized. (When the scan test mode is used to save graphics instead of scanning data, you can disable this option: scanning a graphic, the IRISPen software won't find any text anyway. When the reading mode "Black-and-White Bitmap to Clipboard" or "Greyscale Bitmap to Clipboard" is selected in the IRISPen application window, the option "Recognize" is replaced by the option "Bitmap to Clipboard".)

Enabling the scan test mode and saving the image in a Windows bitmap is just one way of saving small graphics. The IRISPen software offers a much directer way of inserting such graphics in your documents. You don't have to save the image in a file which you insert in your document later on; using the Windows

clipboard, you can paste it directly in your application. On the down side, you cannot retouch the contents of the clipboard!

Select the reading mode "Black-and-White Bitmap to Clipboard" or "Greyscale Bitmap to Clipboard" in the IRISPen application window for direct **image scanning**. (Note that you can no longer select an output target now!) Place the text cursor in your Windows application and scan the graphic. A blinking status window will display the message "Image in Clipboard" for a few seconds to indicate that the image has been copied to the clipboard.

<div align="center">Image in clipboard</div>

Execute a paste operation to insert the image in your document. You can now resize the graphic, move or align it etc. Is there a simpler way of inserting logos, signatures, handwritten annotations etc. in faxes and letters?
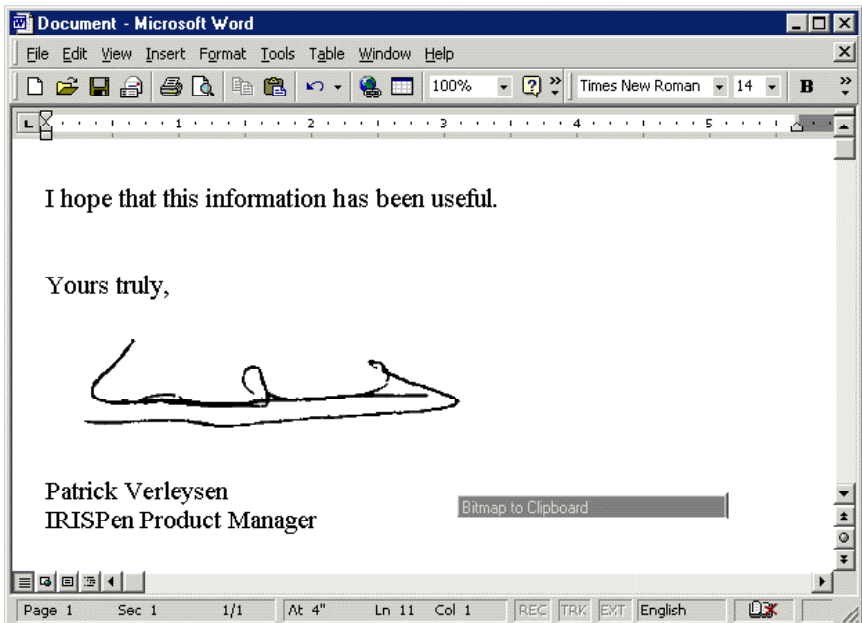
GOOD OCR FOR BAD IMAGES

Back to text! Although we rightly emphasized the importance of steady scanning, it should be noted that the OCR software is very powerful. Given the unstable scanning by the human hand holding a greyscale pen scanner, the recognition technology has to cope with images of very bad quality.
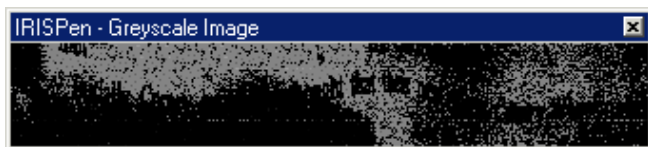
The greyscale image window allows you to see exactly how bad the image quality can be. Colors are a first factor - think of low-contrast documents where the background virtually has the same color as the text. Furthermore, the images can be highly distorted: wavy and skewed lines, contracted, stretched and rotated characters the "raw" material the IRISPen has to process.
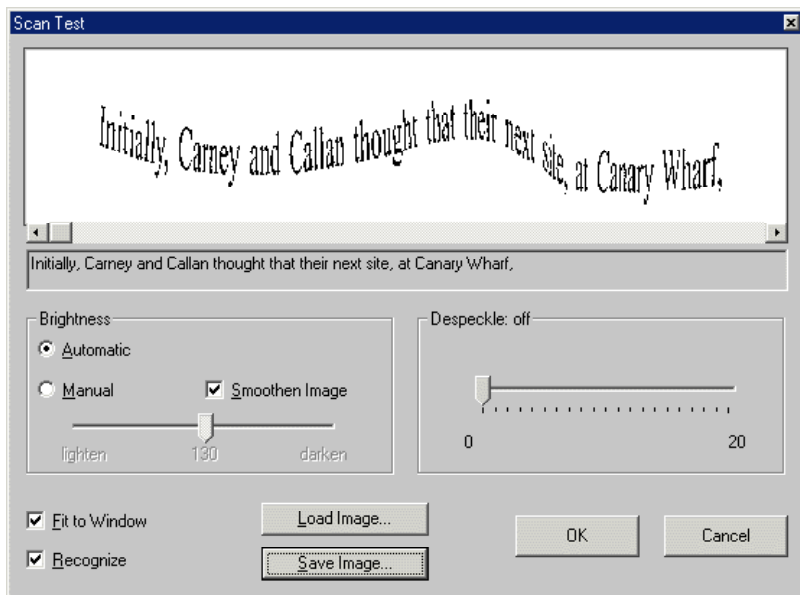
So, the IRISPen offers OCR at its best: the unique **PCR-II technology** - PCR is short for "Pen Character Recognition" - executes character recognition despite every handicap imaginable.

THE IRISPEN IS A LINE SCANNER

The greyscale image and the scan test windows illustrate another character-istic of the IRISPen, **central line tracking**. It is an essential feature: although you scan strokes of about 0.4" (1 cm.) high, containing in many cases several lines of text, the IRISPen recognizes only a single line of text per scan.
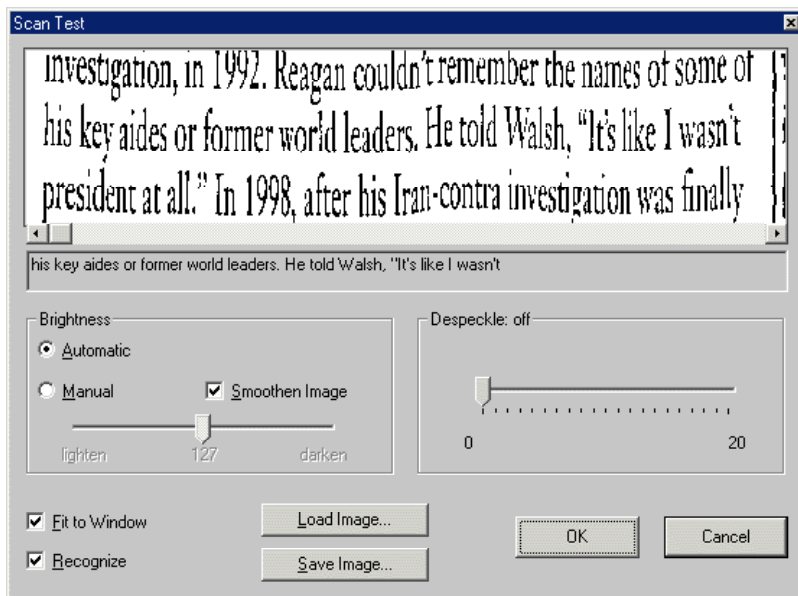
In the above example, we have scanned with a single stroke two complete lines and the major part of a third line. Another column separated by a gutter line

starts on the right side. The IRISPen isolates the correct visual information without any help.

The IRISPen software automatically picks up the "central" line, that is the line directly under or closest to the line guide on the pen scanner. Where the interline spacing is reduced, the line guide on your pen scanner's view window helps you select the correct information.

The scanning height of about 0.4" (1 cm.) explains why you are limited to character sizes of 22 point: bigger characters are simply cut when you scan them. The IRISPen reads **point sizes** of 8 to 22 point (0.11 to 0.31" or 0.28 to 0.78 cm.).
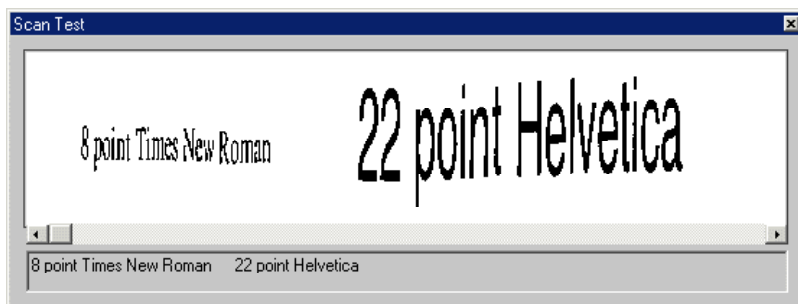
SETTING THE BRIGHTNESS
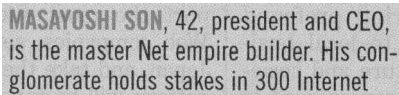
Furthermore, scan tests are useful to set the brightness correctly. As we already mentioned, the scan test window displays a black-and-white version of the scanned image, while the greyscale image window displays the actual scan. We also know that the IRISPen is the first greyscale pen scanner.

In the Executive version, the scan test window is equipped with advanced imaging features to adjust the scanning to the color tones of your document. (You are recommended to keep the greyscale image visible when you're modifying these parameters.)
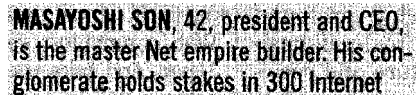
In some instances, scanning in greyscale is mandatory to obtain good OCR results! When text was printed on a color background, scanning in greyscale may create the tone differences that are lacking in black-and-white images. When there is only limited contrast between the text and the background, the background can create "noise" in black-and-white images that renders the recognition difficult or impossible!

Think for instance of black text printed on a dark background: when scanning such a document in black-and-white, you would not be able to "drop" the background color without losing the text information as well.

> MASAYOSHI SON, 42, president and CEO, is the master Net empire builder. His conglomerate holds stakes in 300 Internet

> MASAYOSHI SON, 42, president and CEO, is the master Net empire builder. His conglomerate holds stakes in 300 Internet

Thanks to its powerful intelligent routines, the IRISPen even solves tough cases - here's how our "difficult" image gets binarized!

> MASAYOSHI SON, 42, president and CEO, is the master Net empire builder. His conglomerate holds stakes in 300 Internet

Please note that the IRISPen handles changes - gradual and other - in the background beautifully. To cope with varying colors and varying color intensity in images, the **binarization** is "adaptive". The software chooses the optimal black-and-white threshold *locall*y: a different threshold can apply to each small region of the image.

IRISPen - Greyscale Image

Scan Test

The huge Euronet-operated FTP server contains many popular sites which can be

received at a whopping 2Mbps (megabits per second). The service tells you the specific

The huge Euronet-operated FTP server contains many popular sites which can be

Brightness
- ☉ Automatic
- ☉ Manual ☑ Smoothen Image

lighten   127   darken

Despeckle: off

0   20

☑ Fit to Window   Load Image...
☑ Recognize   Save Image...

OK   Cancel

Should this be necessary, you can optimize the image further for the OCR process. The IRISPen Executive uses some complicated concepts here, and we need to discuss them in detail.

The option "Smoothen Image" renders the greyscale images more homogeneous by "flattening", smoothing out relative differences in intensity. As a result, a sharper contrast is created between the foreground - the text - and the background - a color, artwork etc. We suggest that you leave this option enabled at all times.

The brightness now. By "brightness", we actually mean the black-and-white threshold. As indicated earlier, the setting "Automatic" determines the bilevel threshold automatically but with a great deal of intelligence. Leave the default automatic setting for all normal documents. Apply a different threshold when

necessary by darkening or lightening the black-and-white image: when you darken the image, more pixels become black in the black-and-white version, when you lighten the image, less pixels become black.
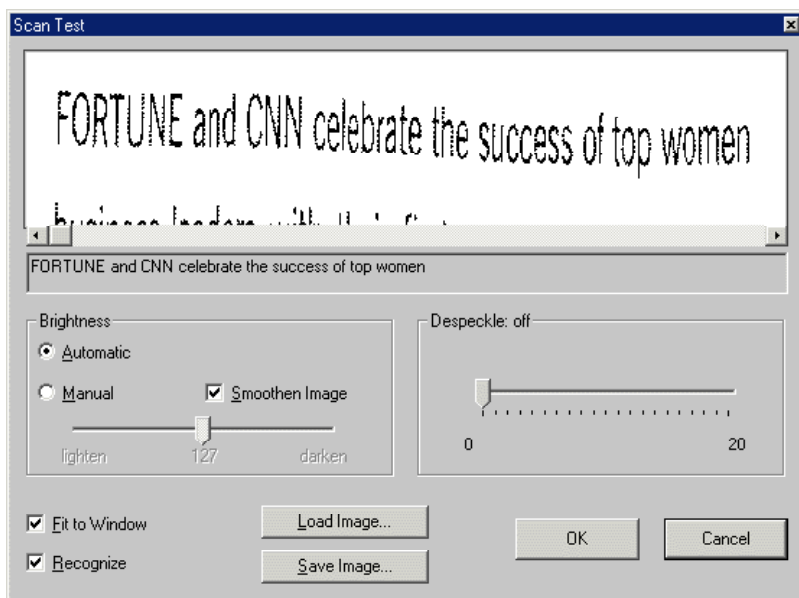
Note that the recognition is re-executed promptly each time you modify the image settings!
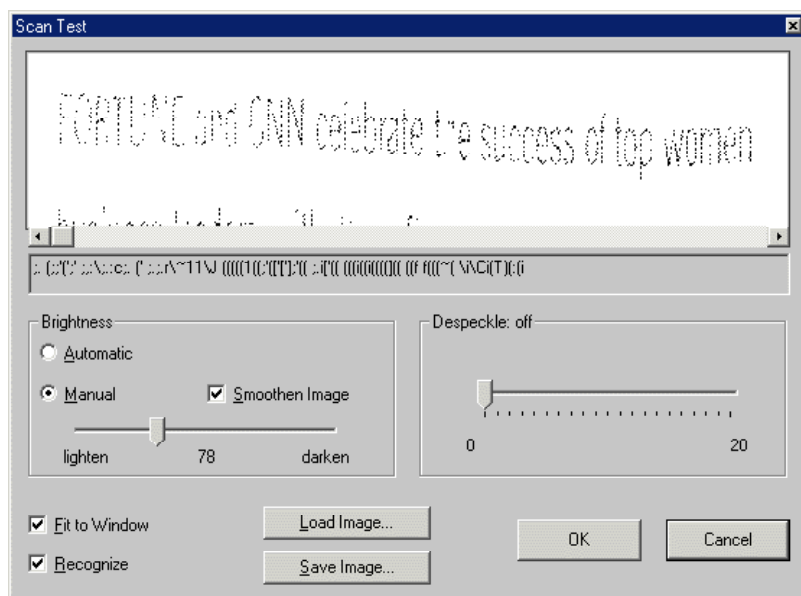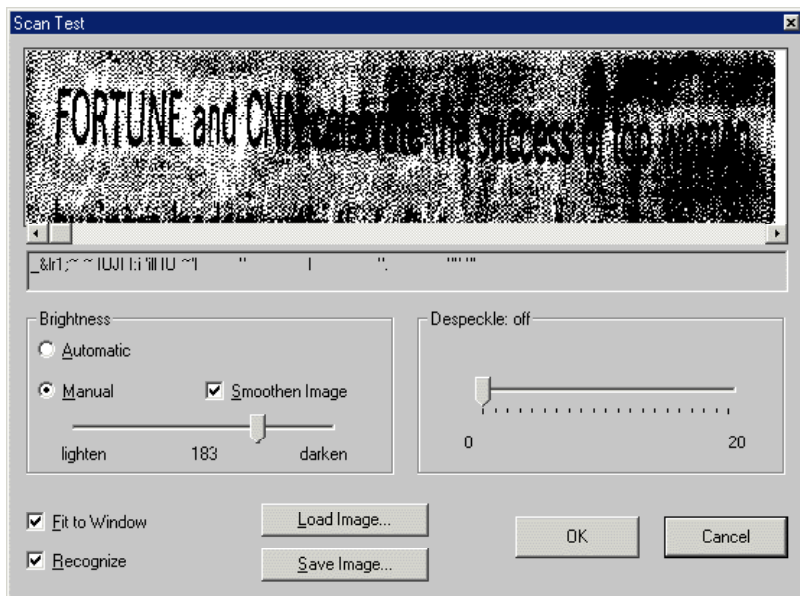
The last option, "Despeckle", concerns the black-and-white image. "Despeckling" means that the "parasite pixels" (also called "salt and pepper noise") will be removed from the black-and-white images.

adapt easily, then   adapt easily, then

Be sure that you don't erase spots that are too big, otherwise you might start erasing the dots on "i", portions of dot matrix letters etc.! (The recognition is re-executed promptly each time you modify this setting.)
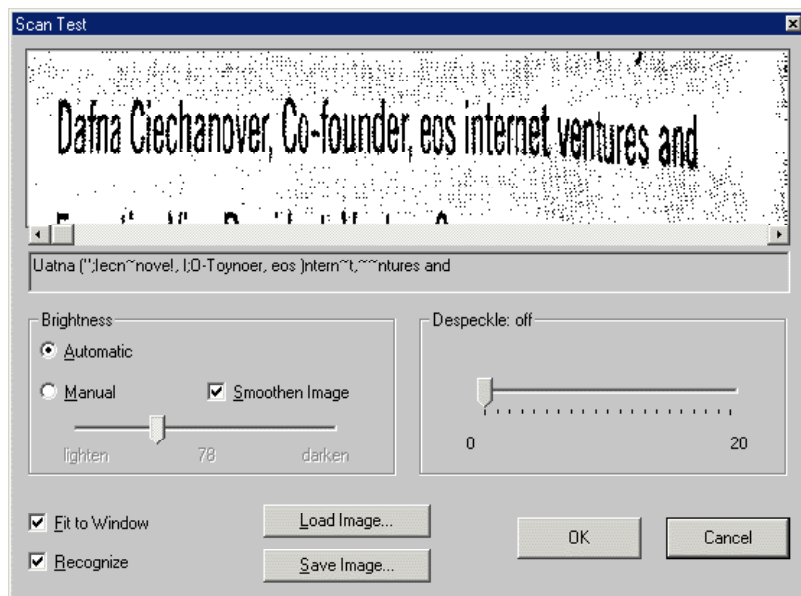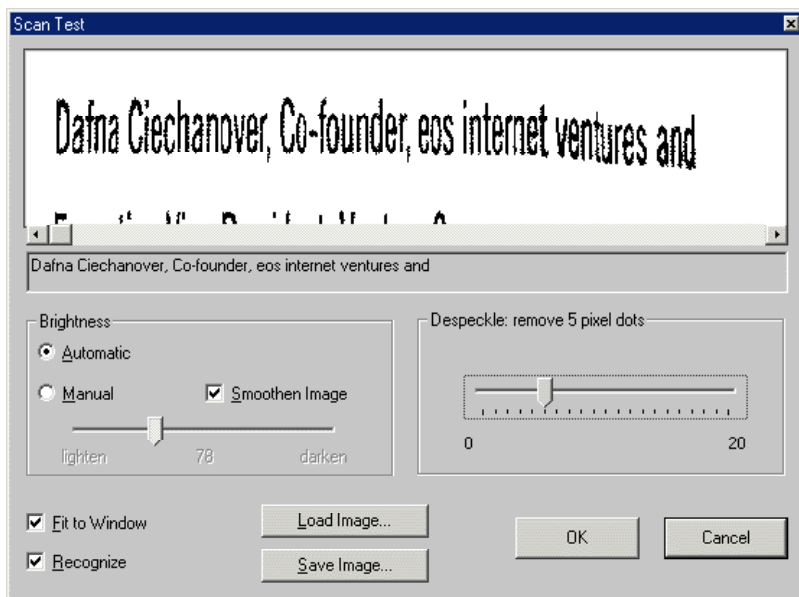
Thanks to the image adjustment features of the IRISPen Executive, a larger range of documents can be read. You make documents legible that couldn't be read in any other way: think for example of documents where the colored background would otherwise render the recognition impossible.

However, a red or yellow background will never be picked up by the pen scanner: red and yellow are the IRISPen's **"drop out"** or invisible **colors**. This implies that all text printed in red and yellow is illegible!

BLACK IS BLACK AND WHITE IS WHITE

Now that we're discussing colors: light text on a dark background *is* legible, provided that you enable the option "Negative Image" under the "Options" menu. (You can also use this option to scan "negative" graphics.)

✔ Negative Image

The image in the greyscale image window is left untouched. The inversion is only visible in the black-and-white image: black "pixels" become white, white pixels become black.

IRISPen - Greyscale Image ☒

FRANCE: FROM 60%
OF EUROPE'S ORGAN

Scan Test ☒

FRANCE: FROM 60%
OF EUROPE'S ORGAN

FRANCE: FROM 60%

The status window indicates whether this option is enabled or not. As light text on a dark background is the small minority, it is disabled by default.

Alphanumeric, English, Application, Negative

HIGHLIGHTING EXCERPTS

Start your wordprocessor now and scan a few text lines. If you have to enter small portions of text into a wordprocessor, the IRISPen is your ideal partner! You can enter several thousands of characters per minute, much more than even the fastest secretary using a keyboard. (Actually, the IRISPen software *processes* about 1,000 characters per second, but no user can scan that fast!)
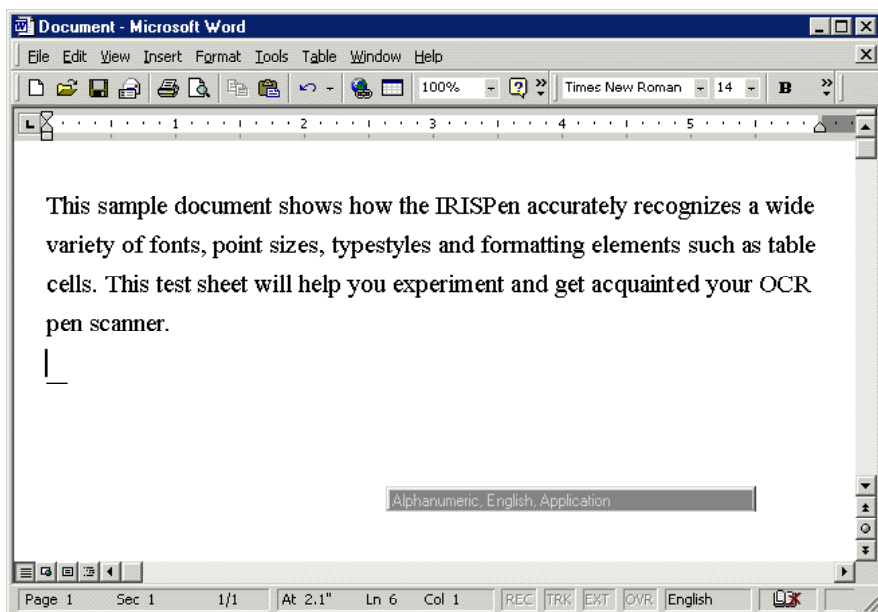
Even a traditional OCR software package with a sheetfed or flatbed scanner is no competition when it comes to scanning **excerpts** and abstracts! To read a few lines of text, you'd have to scan an entire page with your sheetfed or flatbed

scanner, draw a frame around the text snippet to be recognized and wait until the OCR is executed, then transfer the text result to your wordprocessor...

The IRISPen on the other hand offers **real-time recognition**. Slide your pen over your text and in less than a second, it appears on your screen. Use the IRISPen and you can type... with a stroke of a wand! The IRISPen gives you text recognition at your fingertips!

It is important to note that you do not have to wait for the text result to appear on your screen before you start scanning another line.

The IRISPen always inserts the recognized text at the cursor position. The **formatting** as is active in your target application applies. When you have selected, say, Palatino 12 point bold in your wordprocessor, the text you insert will be in Palatino 12 point bold, whatever the original font, typestyle and point size may have been.

READING DOT MATRIX DOCUMENTS

However, in some cases, you have to indicate the font of the source text! The option "Dot Matrix" in the IRISPen application window determines whether you are recognizing "normal" or **dot matrix** printed documents. Dot matrix recognition is limited to the IRISPen Executive.

☑ Dot Matrix

Setting this parameter correctly is mandatory: "draft" or "9 pin" dot matrix symbols are made up of isolated, separate dots, and highly specialized recognition routines are required to recognize them.
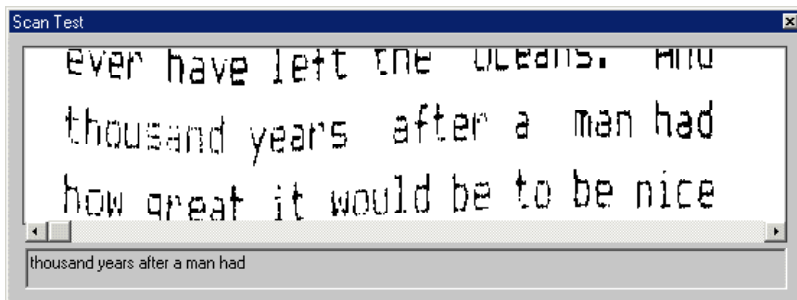
   "Letter quality" dot matrix printing, also called "25 pin" or "NLQ" dot matrix, requires the normal setting, as do the **printing qualities** typeset, typewritten, laser printed and inkjet printed.

   The status window indicates whether the dot matrix reading mode is enabled or not. Obviously, this option is disabled by default.

SCANNING PARAGRAPHS

   When you have a few lines of text, you are recommended to make use of the "Alignment" button on the toolbar (or the "Alignment" command under the "Settings" menu) and specify an **ending code**.

   An ending code is any optional combination of characters and keyboard commands which you add at the end, after the recognized text. To scan in a few lines of text, you'll normally select <Space> as suffix. Doing so ensures that you add a space after each line automatically. If you read two lines, "The second line" and "follows the first.", the resulting output is "The second line follows the first.". The space after the word "line" has been added by the IRISPen software.
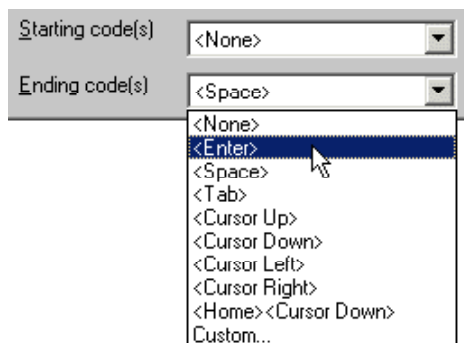
A second option allows you to indicate the start of a new paragraph. The programmable **pen button** takes care of this. Select the "Pen Button" button on the toolbar (or the command "Pen Button" from the "Settings" menu).

Indicate which keyboard command you want to execute or which characters you want to insert when you press the pen button once. To start a new paragraph for example, you will program the pen button to execute the <Enter> key. When you press the button, you execute a carriage return, thus starting a new line.

Pen Button

Single		Double
○	Null	○
●	<Enter>	○
○	<Space>	○
○	<Tab>	●
○	<Cursor Up>	○
○	<Cursor Down>	○
○	<Cursor Left>	○
○	<Cursor Right>	○
○	<Home><Cursor Down>	○
○	Custom...	○
○	Switch reading mode	○

[OK]   [Cancel]
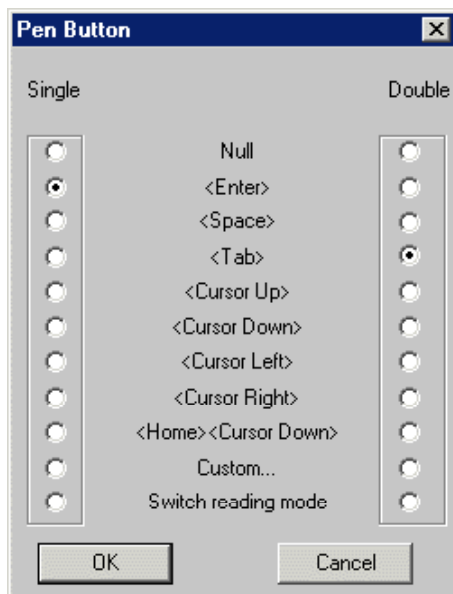
Also note that the pen button makes it easier to insert the scanned text at the right location. When data is to be entered in, say, a database, the <Tab> key can be assigned to the pen button to move the cursor to the next field. (The pen button incorporates some other features which will be discussed later.)

The pen button and **starting** and ending **code** options allow you to execute any combination of characters and keyboard commands: single characters and character strings and one or more keyboard commands are allowed. Even combinations of characters and keyboard commands can be entered automatically without touching the keyboard.

Another option, "Restore Paragraph", also concerns body text. When this option is enabled, hyphens at the end of a line are deleted and the ending code - if one is defined - is ignored.

☑ <u>R</u>estore paragraph

Let's give an example to clear things up. When you read two lines, "The new presi-" and "dent waved from the balcony.", the resulting output is "The new president waved from the balcony.". The end hyphen of "presi-" was ignored and so was the "normal" suffix <Space>. If the two following two lines are: "His lovely wife" and "had joined him.", the output will be "His lovely wife had joined him.", with the space after the word "wife" added by the IRISPen software as the normal ending code.

In other words, this option really does restore **paragraphs** by "recomposing" hyphenated words and adding suffixes only where they are needed.

It is interesting to note that the IRISPen does more than recognize text: keyboard commands you want to execute on your computer system during the scans are taken care of by the IRISPen software. There is no need to lose any time recurring to the keyboard while you are scanning!

ISOLATING TEXT EFFICIENTLY

An additional problem when reading text may be that you scan **clippings** from columnized articles in magazines. How do you isolate the target text comfortably? The "Alignment" command again offers the answer.

┌─ Scanned Area ─────────────────────────────────┐   ┌──────────┐
The quick brown fox jumps over the lazy dog.   │   OK   │
┌─ Output ──└──────────┘
● own fox jumps over the la   (All chars.)
○ own fox jumps over the   (All chars before the last space.)
○   fox jumps over the la   (All chars after the first space.)
○   fox jumps over the   (All chars between the first and last space.)

USER'S GUIDE
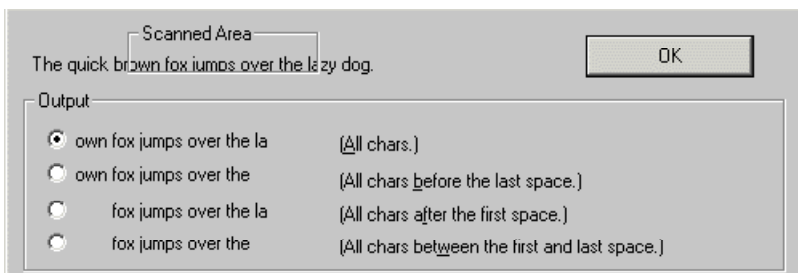
There are four **word cropping** options: you can output all scanned characters, drop the first word (or incomplete word), drop the last word or drop both. By default, all characters are retained.

The last option is particularly useful when scanning closely **columnized text**: to ensure that you do not miss any significant words, you start scanning in the previous column and end scanning in the next column. The IRISPen takes care of deleting the unwanted word parts. In this way, you capture the correct information without having to scan with great precision.

SCANNING MULTIPLE LINES

When you are reading text excerpts, it may be a good idea to enable the option "Multilines" in the IRISPen application window. This option is limited to the IRISPen Executive and the IRISPen Translator.
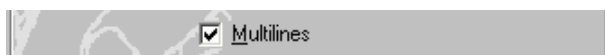
The scanned lines are "buffered" until they are inserted in the target application (or sent to the clipboard). The user presses the pen button to "trigger" the output process. After each scanned line, a blinking status window displays the message "Press pen button to flush" for a few seconds. Only do so at the end of a sentence or paragraph!

Note that you can start rescanning immediately, you can ignore this warning entirely!

EXTRACTING KEY WORDS

An additional cropping function, "Ignore Rejected Characters in First and Last Position", is useful when the IRISPen is used to capture key words.
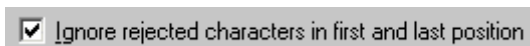
In this case, you will scan with more precision. If you start or stop a bit too early, you have erroneously scanned part of a previous or following word, giving you an unrecognized symbol. This option sees to it that these rejected characters are ignored by the IRISPen. When the gutter between two columns is narrow, or when wordspacing is reduced, this feature comes in handy.

Unrecognized characters are by default represented by a tilde (the "~" symbol), but you can select a different symbol with the command "Unknown Symbol" under the "Settings" menu.

Unknown Symbol

Please specify the symbol for unknown characters

OK

Cancel

~
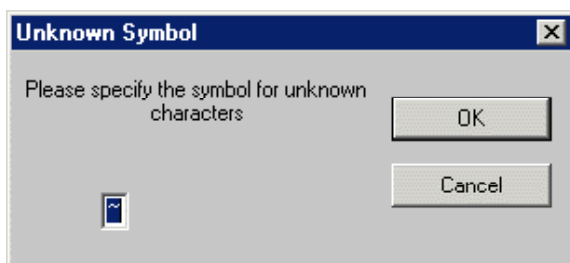
Entering less than a line of text is certainly a major application of the IRISPen. Think of scanning contacts and addresses from brochures, catalogs, business cards in a database, contact or personal manager etc. The IRISPen is the easiest and most practical highlighter imaginable!
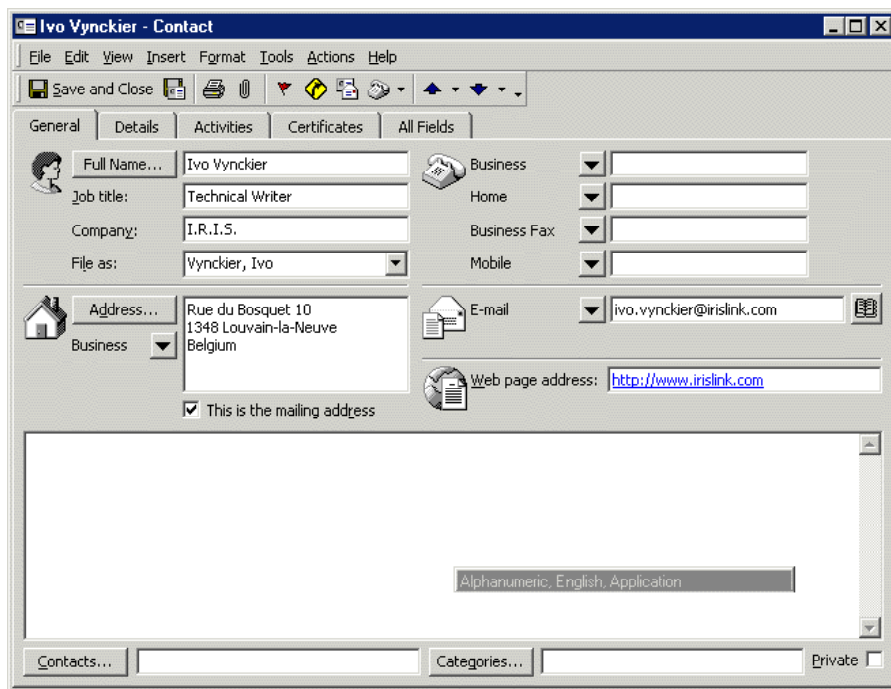
CUT FROM PAPER, PASTE TO WINDOWS!©

Wordprocessors are indeed just a single example of software the IRISPen complements. Actually, the IRISPen will output to *any* Windows application: spreadsheets, databases, office software with e-mail and electronic agendas, personal and contact managers, text editors, encoding utilities, bookkeeping software etc. The list of supported applications is really endless!

The key to the "universality" of the IRISPen is its **keyboard emulation**. The IRISPen effectively replaces your keyboard by imitating keyboard strokes as soon as the application is running in the background, so you can use this tool to insert text wherever you have a text cursor on your Windows PC.

Basically, the IRISPen establishes a "copy and paste" operation: you copy text from paper documents and paste it in your Windows applications.
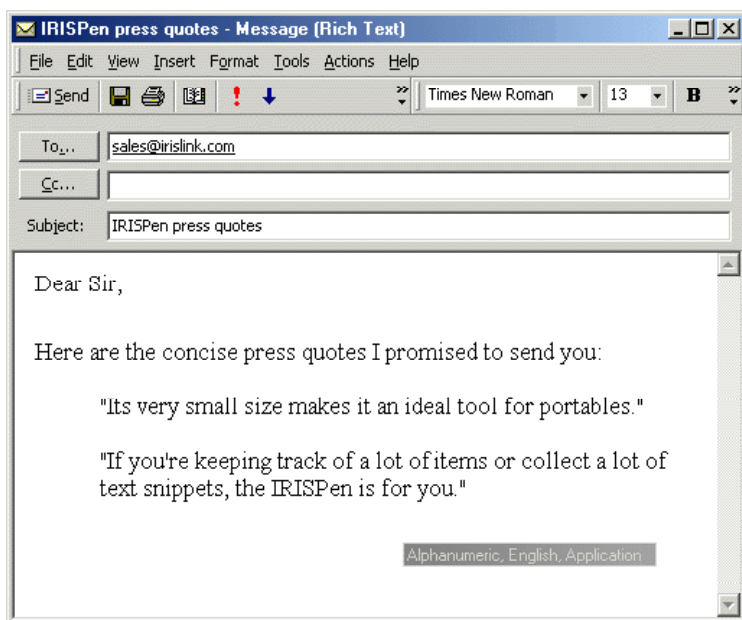
In the above example, we used the IRISPen to enter some quotes from a press article into an **e-mail** message.

COMPLETING DATABASES

When you enter data in a **database**, the possibility to add starting and ending codes at the beginning and/or end of the recognized text is particularly useful.

When you choose for example the keyboard command <Tab> as ending code, the cursor will reposition itself in the next field, ready for the next scan.

Imagine you want to add addresses to your **contact manager**. You scan the company name: the IRISPen puts the company in the corresponding field and relocates the cursor in the name field. You can go on with the person's name, his title, the address, telephone and fax number etc. without having to reposition the cursor each time using the keyboard or mouse.

Here again, a "standard" OCR package falls short of this task as it does not allow you to insert data *directly* in a **database** field. You have to generate a structured text file which you import into your database later on. The IRISPen effortlessly incorporates a flexibility which no other system can offer.

ADD THE CURRENCY... OR DROP IT

Let's discuss some other suggestions concerning prefixes and suffixes: you can use them to add text. A typical application is encoding price lists where you add the currency, for instance "$" or "USD", to the figures you recognize. Or you encode telephone numbers and the system adds "Tel: " in front.

All this means that the IRISPen does more than just recognize text, it also *completes* and *formats* it for you!

The **character conversion** feature is additional proof of this. By character conversion we mean that recognized characters are simultaneously replaced by any combination of characters and keyboard commands: single characters and character string, one or more keyboard command(s) or even combinations of characters and keyboard commands are accepted. Characters can also be deleted (replaced by nothing). A typical application is eliminating the currency in a price list.

Select the command "Character Conversion" under the "Settings" menu. A window shows you a list of all characters; characters to which conversion applies are highlighted. Click on them to see how they are converted. Click the option "No Conversion" to disable the active conversion. The button "Unconvert All" disables all active conversions with a single click.
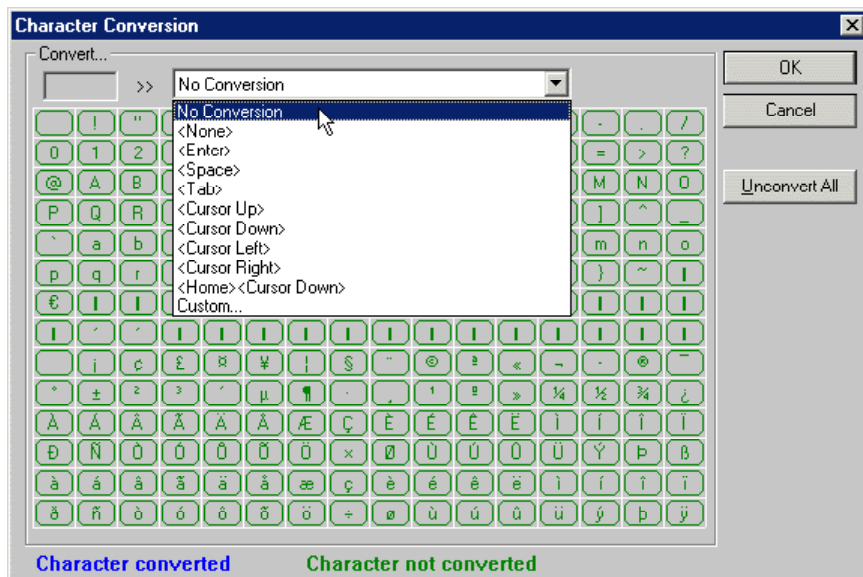
In the example below, the "$" symbol is ignored, erased because the user wants to enter some numeric data in a **spreadsheet**. The figures are relevant, the currency isn't. Scanning "$2.99" leads to "2.99" being entered in the worksheet.
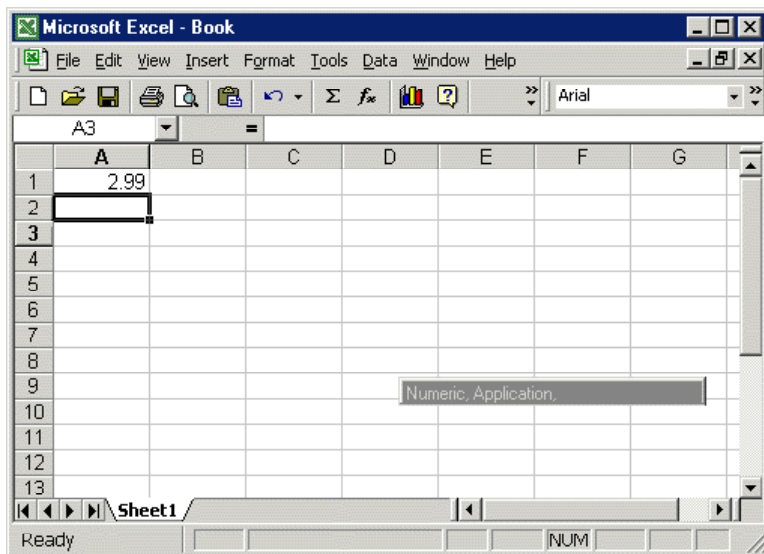
OF ROWS AND COLUMNS

There's another feature of major importance to the scanning of tables of figures: the detection of **table cells** as you find in the example below.

Cells	Column 1	Column 2	Column 3
European	1.462,3	912,10	1,1
American	1,689.37	353.55	3.5

The command "Table Frame Conversion" under the "Settings" menu offers the key. Some common keyboard commands are presented by default, but you

can also enter user defined characters, keyboard commands and any combination of these.

Table Frame Conversion ▶	Null
Character Conversion...	<Enter>
Unknown Symbol...	<Space>
	✔ <Tab>
Bar Codes...	<Cursor Up>
Handprinting...	<Cursor Down>
Output Subset...	<Cursor Left>
	<Cursor Right>
	<Home><Cursor Down>
	Custom...
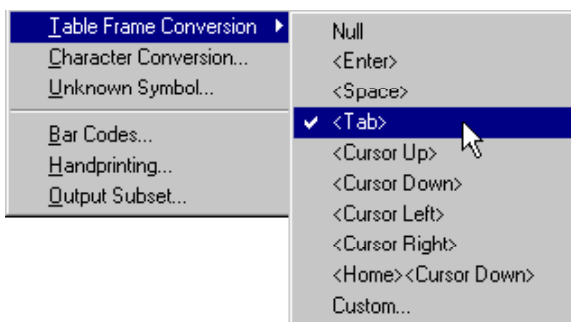
For instance, when you select <Tab>, the IRISPen will execute a <Tab> whenever it encounters a vertical bar. As a <Tab> moves your cursor to the next cell, scanning this table into a spreadsheet will put the individual figures in separate cells. Add <Home><Cursor Down> as ending code to relocate the cursor on the next row.
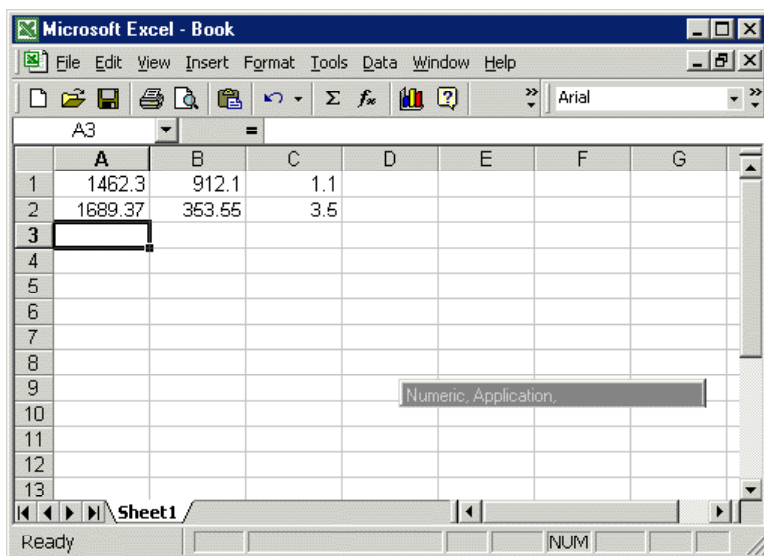
In other words, you can recreate the typical table structure with rows and columns by simply scanning line after line!

Mind you, this only works when continuous vertical bars separate the table columns: the IRISPen software has to encounter a physical mark, detect a border between two columns.

Where the original document does *not* contain vertical bars, the easiest procedure is to quickly add them with a pen and a ruler. Another solution is reading the figures one by one and programming the pen button with the "Pen Button" button on the toolbar. The single pen button click should be programmed as <Tab> - the cursor is moved to the next cell - , the double-click as <Home><Cursor Down> to start a new row.

In some cases, even the command "Character Conversion" from the "Settings" menu can offer the answer as it allows you to replace scanned characters

by a cursor command. Here's an example: to insert the line "2,55£   5.99£   65£" in a worksheet, it suffices to replace the symbol "£" by <Tab> to place the individual figures in separate cells.

MORE FUN FOR NUMBER CRUNCHERS

Entering tabular data in a spreadsheet is one of the most popular IRISPen applications. Actually, the IRISPen is capable of entering numbers faster and more accurately than any human being can. It's a text book example of "machine beats man"!

The IRISPen will recognize your tables of figures with even more precision when you set the **reading mode** to numeric. This is accomplished in the IRISPen application window. Doing so limits the set of recognized characters to **numeric symbols**, in which case you will only recognize the characters '0' to '9' and some punctuation symbols such as ".", ",", some currencies - the dollar and Euro sign etc. By doing this, you eliminate any chance of recognition errors such as the confusion between '5' and "S", '0' and "O", '9' and "g", '8' and "B" etc.

In fact, you can fully customize the set of recognized characters for your encoding utilities by creating **character subsets** that exclusively contain those symbols you want to recognize. (Each time you do, the recognition speed is increased somewhat.)

An example: you want to read figures from a price list. Unfortunately, the price list also contains some currency symbols such as $ and £. The standard numeric character mode obviously won't do. You can nevertheless read such data faultlessly by extending the basic numeric symbol set with the symbols "$" and "£".

Click on the "Output Subset" button on the toolbar (or select the command "Output Subset" under the "Settings" menu).

Click on a symbol to select and deselect it. The buttons "Select A-Z", "Select 0-9" etc. provide a faster method of selecting ranges of symbols; click on these buttons to add and remove symbol ranges.
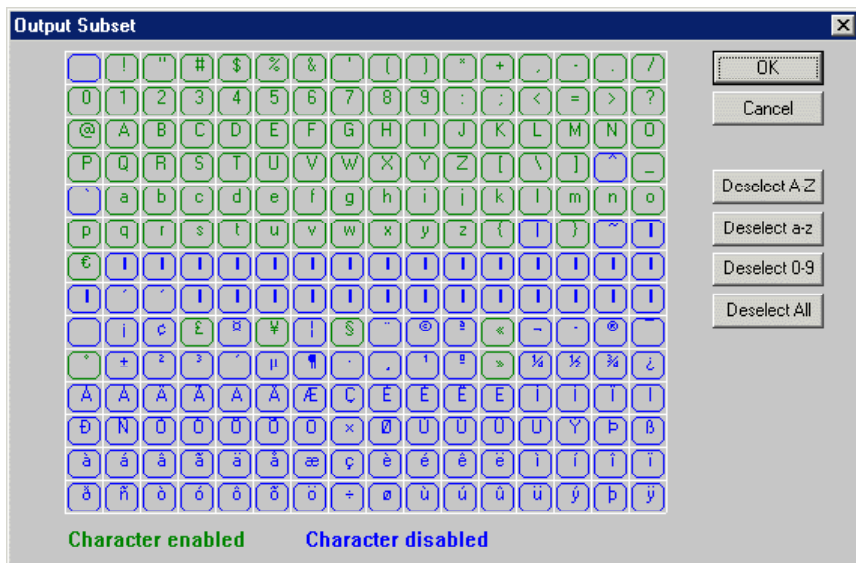
To read standardized financial documents such as printed checks, bonds, bank statements etc., you could limit recognition to the **OCR-B** characters by enabling exclusively the numeric symbols, "<", ">" and "+". (OCR-B is actually more than a character set. It is also a highly specialized typeface used only in standardized documents, and I.R.I.S. has a specialized version of the IRISPen software for the financial market.)

Scan Test

3694003941838731600042522+0000104777+145057573195<

`3694003941838731600042522+0000104777+145057573195<`
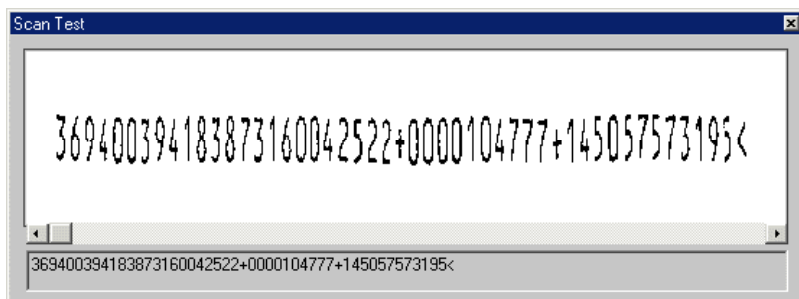
A third example is reading data strings with a specific syntax. Think for example of the serial numbers found on product labels. Such character strings go beyond the numeric symbols, but never contain the full alphabetic range of characters.

The reason for limiting the character set is that you keep excluding recognition errors as the confusion between '5' and "S", '0' and "O" etc. and that you make the recognition somewhat faster.

However, the IRISPen does not recognize *all* characters which are displayed in the output subset window! Some special characters, like © and ®, are not recognized by the IRISPen but can be used in character and table frame conversions, in starting and ending codes and be inserted with pen button clicks.

The window below tells you exactly which symbols can and cannot be recognized - the highlighted symbols are *not* recognized by the IRISPen.
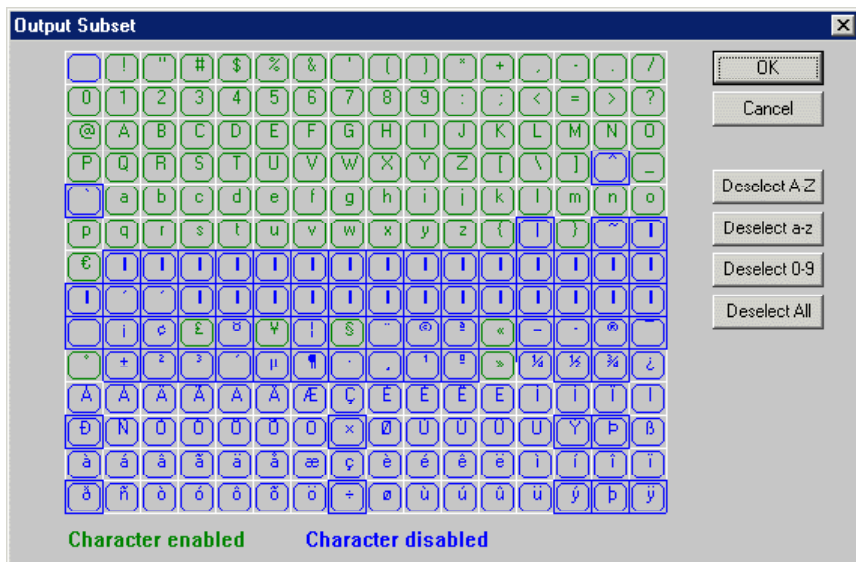
THE READING PEN... EVEN SPEAKS!

When you are scanning figures and structured character strings, precision and speed are of major importance. Thanks to the IRISPen's integrated "Read & Speak" speech technology, you can go ahead and scan your text without looking up to the screen to verify the result.

"Hands- and eyes-busy" data entry become possible as the **speech synthesis** provides you with clear, natural sounding auditive feedback of all recognized text. Encode addresses and your sound card will pronounce them, enter figures into a spreadsheet or encoding utility, enter invoices into your bookkeeping software and the IRISPen software reads the figures aloud for verification. Sec-

ondly, the speech feature opens a world of **multimedia** applications to the IRISPen users.

Refer to chapter 3 to learn more about it all!

THE IRISPEN TAKES YOU AROUND THE WORLD

Customizing the character set is actually also what you do when you select a language in the IRISPen application window.

The IRISPen is far from limited to English: up to 55 **languages** are supported! All American and European languages are supported, including the Central-European languages, Greek, Turkish, the Cyrillic ("Russian") and the Baltic languages.
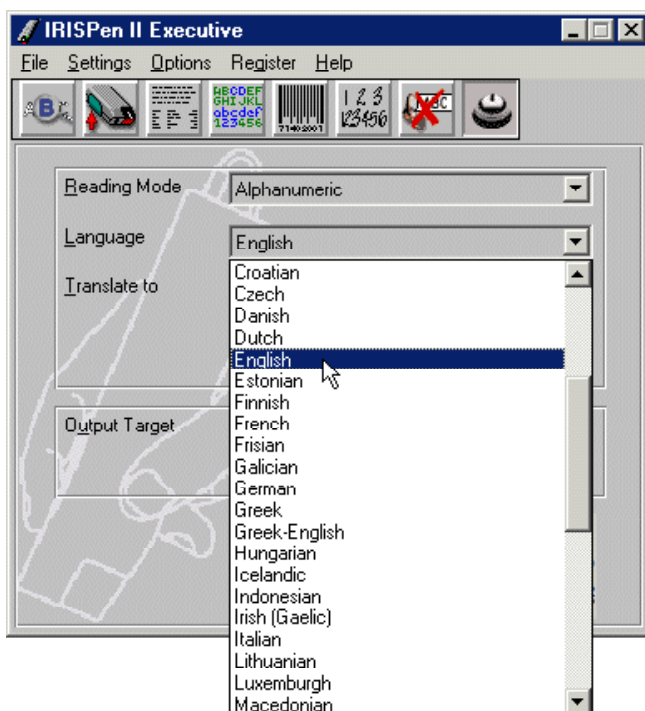
Also note that the British and American - or should we say "international"? - variants of the English language are distinguished. The same goes for Spanish and Mexican.

You can press a letter key to move to it directly: if English is currently selected, and you want to select Hungarian, you can click the "H" key on your keyboard to go directly to the Hungarian language. When several languages have the same initial, press the letter several times to go through the options. Let's give an example: the IRISPen reads Polish and Portuguese. By pressing "P" once,

you select Polish, by pressing "P" a second time, you select Portuguese, and by pressing "P" a third time, you're back on Polish. (To go to *another* letter, say T, press BackSpace before you enter the "T" character.)

It takes the appropriate Windows configuration to display Central-European, Greek, Turkish, Cyrillic and Baltic characters. You may have to install the Windows module "Multilanguage Support" before your Windows system is able to cope with these languages. Select the icon "Add/Remove Programs" under the "Control Panel" to find out if this Windows module is installed on your PC.

USER'S GUIDE
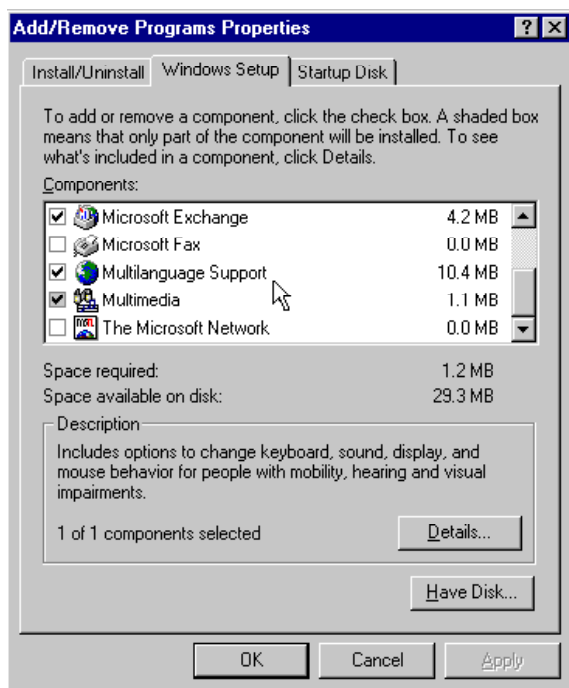
To read exotic languages, you also need to enable the correct keyboard layout under Windows! By doing so, an appropriate font is activated.
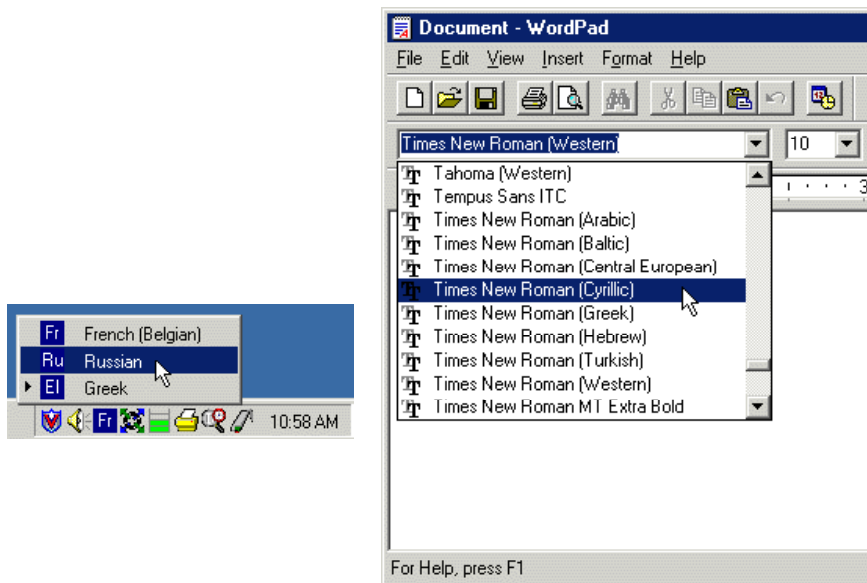
The Windows control panel "Keyboard" allows you to install the necessary keyboard layouts.
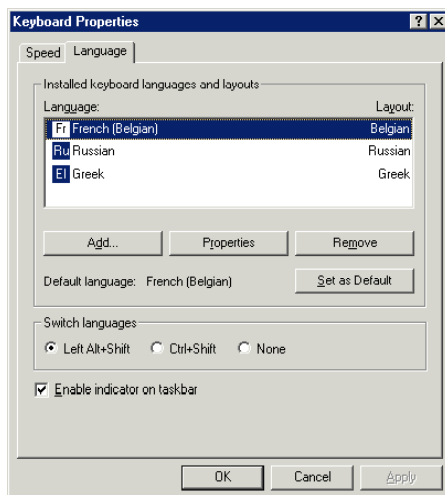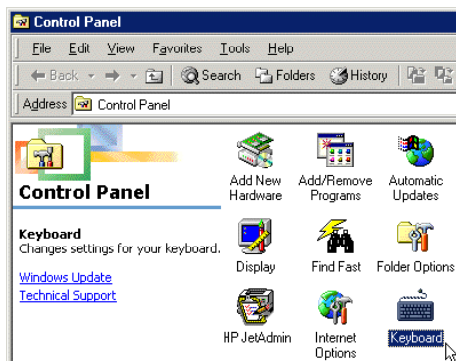
Selecting the proper document language is imperative. Based on the selection of a language, the software knows which **symbol set** to recognize. Multilinguistic support ensures that "exotic" characters such as ç, ß, ñ, γ and ø are recognized correctly.

Secondly, the software extensively uses **linguistic databases** to validate its results. Suppose that you have to read the word "president" where an ink stain makes the "r" look like an "f". Looking things up in the English lexicon, the IRISPen will detect autonomously that the word "president" is being read and that it doesn't make any sense to recognize the symbol "f". This **"self-learning" technique** is of course highly dependant on the linguistic context.

Linguistics also offer useful help to solve **ambiguous cases** such as an "O" which might be mistaken for a '0'. Another typical example is the letter "l" and the number '1' which have an identical form in many fonts - think of texts produced on old typewriters! The linguistic context helps to determine whether you are dealing with "l" or '1'.

The illustration below shows various shapes of '1' and "l". The shapes on the first line are unambiguous, the shapes on the second line are ambiguous, but linguistics can solve them.

THE IRISPEN CHANGES LANGUAGES AS NEEDED

But the buck doesn't stop here: the IRISPen can switch languages in the middle of a sentence without any help from the user! When Western words pop up in Greek and Cyrillic documents - many untranscrible proper names, brand names etc. are written using the familiar Western symbols -, the IRISPen can switch to the correct alphabet automatically. In other words, you can activate a **mixed alphabet** of Greek or Cyrillic and Western characters.

Be sure to select "Greek-English" or the appropriate Cyrillic language setting - for instance "Byelorussian-English". In other words: don't try to just select "Greek" or "Byelorussian" as document language and hope that the Western symbols will come out fine!

Here's an example where a Greek text contains an English word!
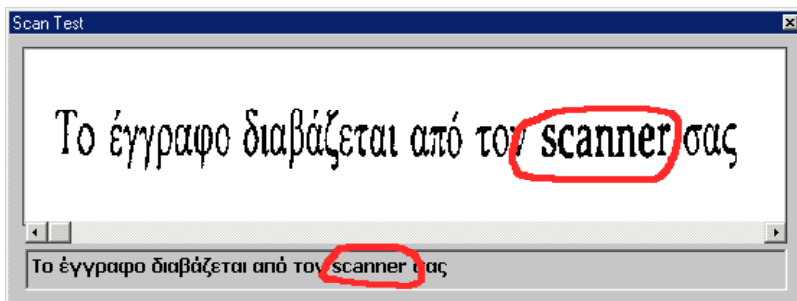
The result looks like this when opened with the wordprocessor - don't forget to enable the "exotic" keyboard layout to read and display such texts.
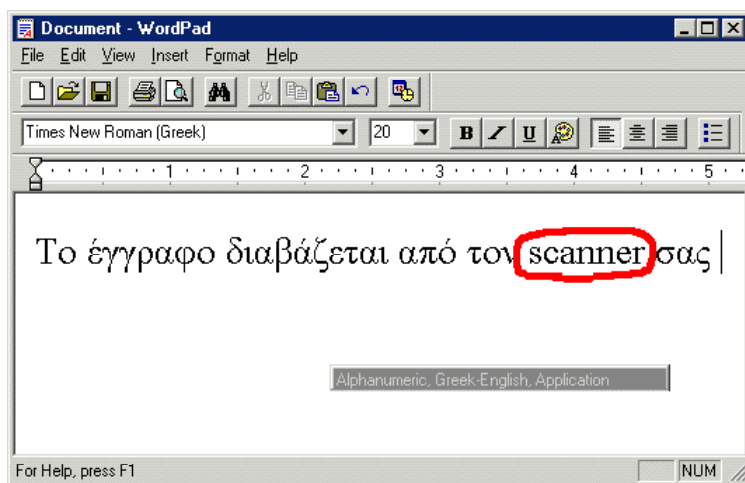
To **mix other languages**, simply select the language with the most extended character set. If you have a document where the, say, French translation is placed alongside an English text, you have to select French as language to ensure that the accentuated characters such as ç, é and ù get recognized correctly.

UNITY IS STRENGTH

So much for the major features and options which enhance and customize the operation of the IRISPen. One element should still be stressed here: the fact that these commands are by no means separate elements! On the contrary, for optimal use of the IRISPen, you will typically combine several features.

Let's give two examples. To scan an English columnized **text** into a wordprocessor, you could do the following: select English as the document lan-

guage in the IRISPen application window, enable a cropping option to read the middle column comfortably, have the paragraphs restored and add a space as ending code with the "Alignment" command and finally program the single pen button click to execute <Enter> and start a new paragraph.

To scan a table of **figures** in a spreadsheet, you could combine the following features: limit recognition to the numeric characters in the IRISPen application window or enable a custom character set with the "Output Subset" command, filter the currency "$" from the output with the "Character Conversion" command, execute a <Tab> when vertical bars are met with the "Table Frame Conversion" command to put the scanned figures in separate cells and program <Home><Cursor Down> as ending code with the "Alignment" command to start a new row automatically.

THE LAW OF INERTIA

Correctly setting these operational parameters doesn't take much time. Still, the IRISPen software allows you to save them in settings files. This is accomplished in two ways.

First, you can save settings as default configuration by enabling the option "Save Settings on Exit" under the "Options" menu.

✔ Save Settings on Exit

When you quit the application, the current settings will be saved in the file DEFAULT.IPC. (When this file is erased, the factory settings apply.) These settings will take effect as soon as you start the software up again.

Secondly, you can save configuration files for specific applications. The command "Save Settings" from the "File" menu saves the current IRISPen settings for future use.
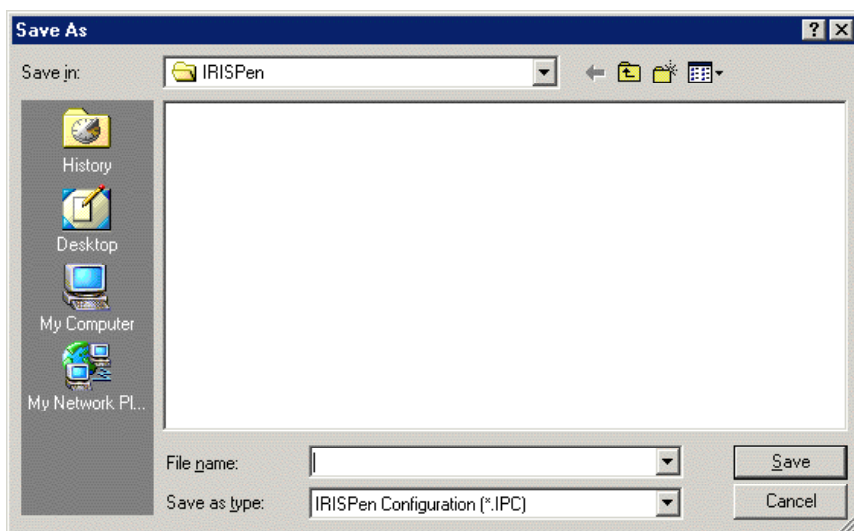
For clarity, give meaningful names to your IRISPen configurations, for instance GERMAN, TABLE etc.

The command "Open Settings" under the "File" menu loads IRISPen configurations into memory.

Settings files contain *all* operational parameters: the reading mode, language and output target, the programming of the pen button clicks and hotkeys, table frame and character conversions, the output subset, alignment options and starting and ending codes, the presence and position of the greyscale image window and whether bar code reading functions as secondary reading mode, whether the dot matrix reading mode, the left-handed scanning, the image inversion and the speech feature or the system sound are enabled.

Also note that the last four settings are directly available on the "File" menu; simply click a configuration to load it.
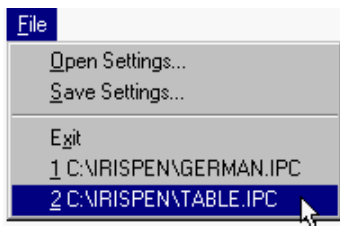
RESTORING THE APPLICATION TO THE FOREGROUND

As we're talking ease of use now, we'll go on with some additional features. We already know that the IRISPen software is equipped with a **toolbar** for quick access to all frequently used commands. Let's briefly repeat what the buttons stand for.
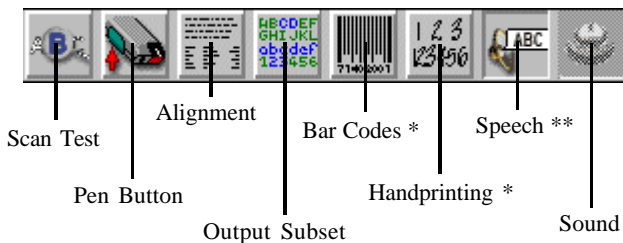
* Limited to the IRISPen Executive
** Limited to the IRISPen Executive and the IRISPen Translator

What else is there? What is, say, a quick way of accessing the IRISPen application window, for instance to modify the language, output subset etc. First of all, you can restore the IRISPen application by right-clicking the IRISPen control on the Windows taskbar and selecting the command "Open".
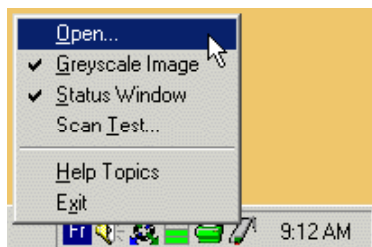
To switch *directly* to the IRISPen application, you can right-click or double-click the status window. The status window does more than provide information, it also gives you direct access to the IRISPen application!

Finally, you can define a **hotkey** with the command "Hotkeys" under the "Settings" menu. Enter a key combination to restore the application. Evidently, you should not use key combinations such as Ctrl+V which are used by your target application(s).
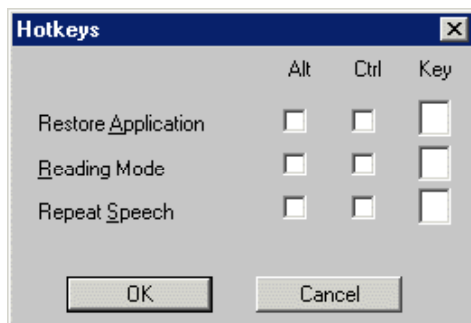
So, **restoring** the IRISPen **application** which normally runs in the background is a cinch. Another interesting feature is that you can switch between the reading modes with equal comfort. The command "Hotkeys" under the "Settings" menu again offers the solution.

USER'S GUIDE

When you have selected "Alphanumeric" as your reading mode and you execute the hotk, the IRISPen switches to the mode "Bar Codes". When you repeat the hotkey a few times, you switch to "Black-and-White Bitmap to Clipboard", "Greyscale Bitmap to Clipboard", "Handprinting", "Numeric" and back to "Alphanumeric".

MORE ABOUT THE PEN BUTTON

And why not use the pen button to switch modes? You can program a button click to **switch between** the **reading modes**. To encode documents where text and figures are mixed on the page, just clicking the pen button can do the job.
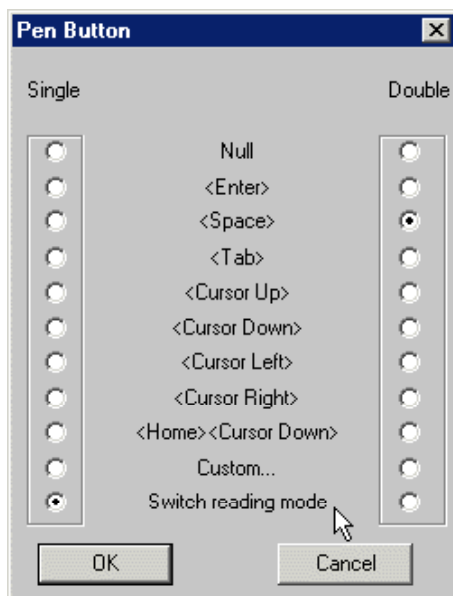
The fact that the single and double click of the pen scanner button are programmed separately adds all the more flexibility! Let's give some illustrations. Imagine you are encoding addresses from business cards, brochures, catalogs etc. you collected during a trade show in a database. Depending on the addresses, some fields will be skipped. For some addresses, the country will be added, and for others, it won't, some contacts have an e-mail address while others don't etc.

To encode these data efficiently, the single pen button click is programmed as <Tab>, even though we already have <Tab> as ending code. The suffix <Tab> moves the cursor to the next field automatically, and you click the pen button once when a database field should be skipped. To start a new record when an address was completed, you normally click a button with your mouse. Here, you can program the double button click of the pen button to correspond with that button's shortcut.

In this example, we can skip the empty fields with a single button click and start a new record by programming the double pen button click as Alt+S, the shortcut for the "Save and Close" button.

Or you have to encode forms with addresses and coded numbers in a database. The codes are strictly numeric, the address is of course composed of alphanumeric data. To add complexity, the data must be distributed among different database fields. Here's how you can proceed: use the single button click to move the cursor to a new database field and use the double button click to toggle between the alphanumeric and numeric reading mode.

Spreadsheets offer another application when figures are read one by one: the single click is programmed as <Tab> and allow to move the cursor to the next cell, the double-click is programmed as <Home><Cursor Down> and moves the cursor to the next row.

GETTING ON-LINE HELP

This concludes our overview of the IRISPen. Some last-minute information may not be included in this manual. We thus recommend you to consult the on-line help system for additional information on the IRISPen's capabilities.

Go to the "Help" menu to do so. The command "Help Topics" and its shortcut key F1 allow you to navigate through the many help topics.

USER'S GUIDE
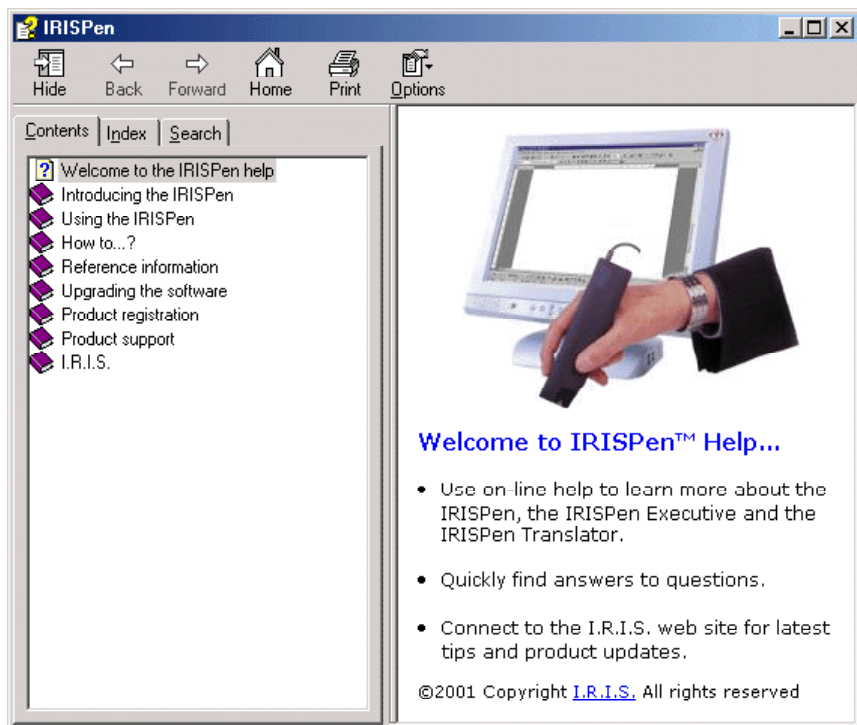
ACQUIRING UPGRADES

The command "Upgrade" under the "Register" menu tells you how to upgrade the IRISPen software to the multifunctional reading pen IRISPen Executive or the translation and language learning tool IRISPen Translator, and how to acquire extra options for your IRISPen.

Taking the Upgrade Path

It's a matter of modules!

Basically, it's a modular approach: by installing the **software-only upgrades**, you add whatever extra feature you don't have yet.

You can upgrade from the IRISPen Standard to the IRISPen Executive and Translator, from the IRISPen Translator to the IRISPen Executive and from an IRISPen Executive to an IRISPen Translator.

Migrating to an IRISPen Executive

- From an **IRISPen Standard**: the Executive software adds four reading capabilities – speech synthesis (in one language), bar code reading, handprinting recognition and the recognition of dot matrix documents. Also offers multiline reading and image adjustment.

- From an **IRISPen Translator**: the Executive software adds three capabilities – bar code reading, handprinting recognition and recognition of dot matrix documents. Also offers image adjustment.

Basically, the IRISPen product range uses a modular approach: by installing an upgrade, you add whatever extra feature you don't have yet. You can upgrade from the IRISPen Standard to the IRISPen Executive and to the IRISPen Translator, from the IRISPen Translator to the IRISPen Executive, from an IRISPen Executive to an IRISPen Translator etc.

MORE ABOUT THE IRISPEN EXECUTIVE...

The software-only upgrade to the **IRISPen Executive** offers numerous major add-on capabilities: speech synthesis, bar code reading, handprinting recognition, recognition of dot matrix documents, advanced image adjustment features and multiline scanning. As a result, your IRISPen becomes unquestionably the most versatile scanner available!

We already mentioned that the **speech synthesis** capability is limited to the IRISPen Executive - where the user is entitled to a single language of his choice - and to the IRISPen Translator - which is equipped with the speech libraries of the two languages in the language pair. (Speech modules for additional languages can further extend his system. Ten languages are currently available: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.)

Bar code reading opens a world of new applications to IRISPen owners as their pen scanner becomes an ergonomic and compact bar code scanner for the office environment. Information desks, marketing surveys, customer identification, order picking and many other examples of high accuracy data entry in the office environment can benefit greatly from this bar code reading solution.

All widespread bar code standards or "symbologies" are supported: Code 39, Code 93, Codabar, EAN-8, EAN-13, UPC, Interleaved 2 of 5 and MSI Pharma. (Support of the Code 128 symbology is optional.) The user can but does not have

to specify the symbology: the user-friendliness is such that the IRISPen automatically detects which bar code standard has been scanned! There is no programming or parameter setting involved in its operation, its use is automatic and transparent.

Bar Codes

☑ EAN-13	☑ Verify and Erase Check Digit
☑ EAN-8	☑ Verify and Erase Check Digit
☑ UPC-A	☑ Verify and Erase Check Digit
☑ Codabar	
☑ Interleaved 2 of 5	
☑ Code 39	☐ Full ASCII Set
☑ Code 93	
☐ Code 128	
☐ MSI Pharma	

OK   Cancel   Default

You can even enable a **mixed reading mode** where bar codes are read alongside other data - for instance "normal" text. The IRISPen "auto-detects" which type of data is being read. For example, you can scan address labels where the printed data is complemented by a bar code without ever changing the reading mode!

Reading Mode	Alphanumeric	▼
Language	English	▼
Translate to	<None>	▼
	☑ Bar Codes	
	☐ Dot Matrix	

Encoding forms and labels can also benefit greatly from numeric **handprinting recognition**. This proprietary recognition technology is based on more than 1 million writing samples! All natural writing styles - American and European - are supported; contrary to most other handprinting recognition solutions, no training phase or imposed style is required.

Scan Test	☒

310 - 1033310 - 55

310-1033310-55
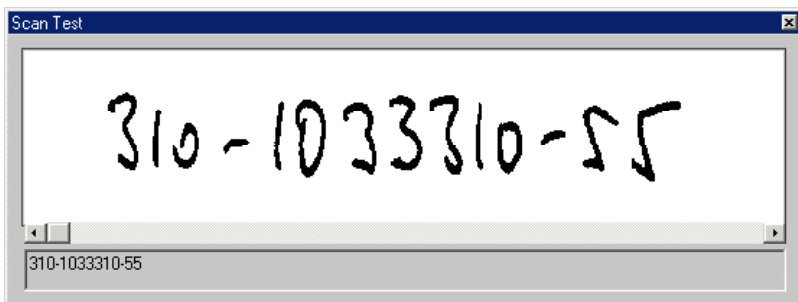
Furthermore, the IRISPen Executive supports **dot matrix recognition**. Dot matrix symbols are made up of isolated, separate dots, and highly specialized recognition routines are required to recognize them!

Scan Test

ever have left the oceans. And

thousand years after a man had

how great it would be to be nice

thousand years after a man had
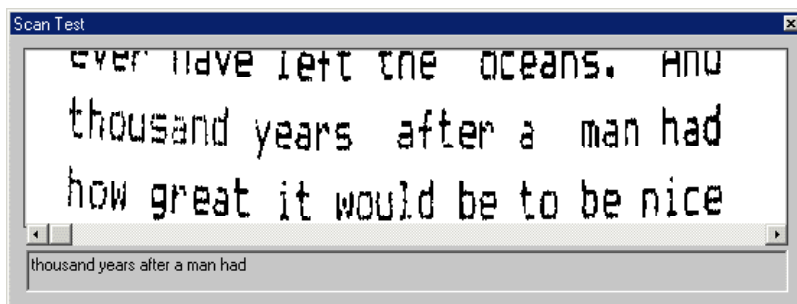
Multiline scanning increases the "user-friendliness" of the IRISPen: several lines are "buffered" before they get inserted in the target application. You can send full paragraphs consisting of several lines to your Windows applications at once. The multiline reading mode also leads to a more natural prosody: when the text-to-speech module has to pronounce individual lines, rather than full sentences, the intonation may be somewhat less than "human".

Finally, the IRISPen Executive is equipped with advanced **imaging features** to adjust the scanning to the color tones of the documents to be scanned. As you can optimize the images generated by the pen scanner for the OCR process, a much larger range of documents becomes legible. Think for example of documents where the colored background would otherwise render the recognition impossible.

... AND THE **IRISPEN TRANSLATOR**

The software-only upgrade to the **IRISPen Translator** offers the following major add-on capabilities: speech synthesis in two languages, bidirectional sentence translation of a language pair and multiline scanning.

The IRISPen Translator is a great tool for language learning and translation aid: you encode and translate foreign language texts by scanning them, and you hear the words spoken.

The IRISPen Translator provides bidirectional, real-time **sentence transla-tion** of all scanned text; the text is inserted in your active application, the transla-tion "pops up" on your computer screen in a speech bubble.
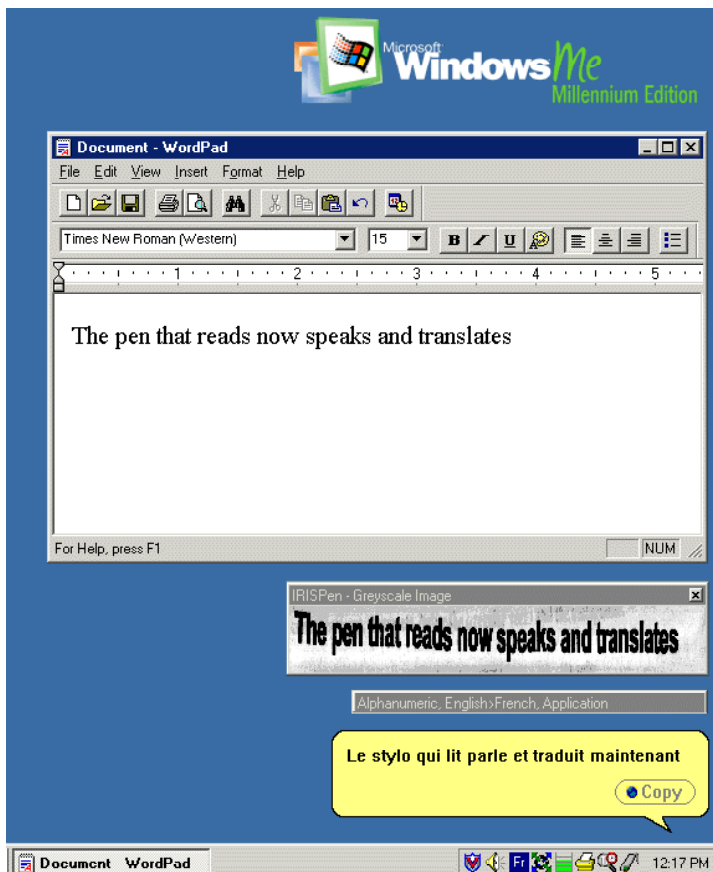
The IRISPen Translator software also offers **clipboard translation**. You translate texts without scanning, the text to be translated is simply copied to the clipboard: think of texts entered on the keyboard, wordprocessor files you open, e-mail messages you receive, Internet web pages you consult etc.

The IRISPen Translator English-French translates English into French and vice versa. Extra translation libraries can be installed. The currently available versions are: English-French, English-German, English-Italian, English-Portuguese and English-Spanish. Contact I.R.I.S. to learn the availability of new language pairs.

The **speech synthesis** opens the door to a world of new applications! To learn how to speak a language, you can just scan a text and hear the PC pronounce the words. The speech feature is offered for the language pair: the English-French version of the IRISPen Translator for instance has the English and French speech libraries. (Speech modules for additional languages can further extend the user's system. Ten languages are currently available: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.)

The **multiline scanning** is mandatory for the IRISPen Translator. As sentences in most cases comprise several lines of text, "multiline" scanning allows to buffer several lines before the output process - text insertion, pronunciation and translation - is triggered. Otherwise you would pronounce partial sentences and (try to) translate ungrammatical text lines as you scanned them, with unavoidably weak results.

EQUIPPING YOUR IRISPEN WITH OPTIONS

Finally, there are some options available for the IRISPen product range.

The **speech synthesis** of the IRISPen Executive is limited to a single language of your choice, the IRISPen Translator is equipped with the speech libraries of the two languages in the language pair. You can extend your system by installing text-to-speech modules for additional languages. Ten languages are cur-

rently available: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.

Secondly, the IRISPen Executive supports all widespread bar code symbologies, but support of **Code 128**, a specialized bar code used in the financial world, is optional.

The same goes for the recognition of the specialized **banking fonts** (OCR-A, OCR-B, E13B and CMC-7). This reading capability is offered by the special software version for the financial world. Contact I.R.I.S. for more information.
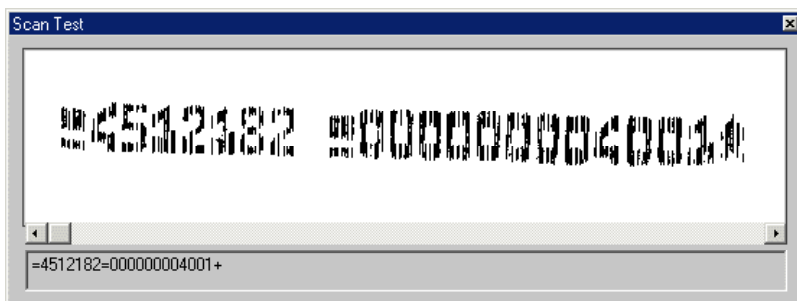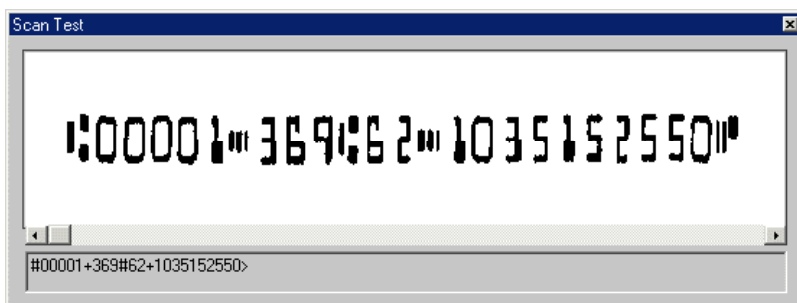
Chapter 3
READ & SPEAK

If you feel the IRISPen should not only read but also talk, your problems are over! The "Read & Speak" speech synthesis as integrated in the IRISPen Executive and the IRISPen Translator brings spectacular new applications to all IRISPen owners.

Where the IRISPen is by definition a "hands- and eyes-busy" data entry device, **text-to-speech** provides clear, natural sounding auditive feedback of all recognized data.

When you are scanning bar codes, figures and structured character strings such as bank statements, telephone and social security numbers, precision and speed are of major importance. Text-to-speech reduces the need to read information on the computer screen to a minimum, you can go ahead and scan without hardly looking up to the screen to verify the result. Encode addresses and your sound card will pronounce them, enter figures into a spreadsheet or encoding utility, enter invoices into your bookkeeping software and the IRISPen Executive software reads the figures aloud for verification.

Secondly, the speech feature opens a world of **multimedia** applications to the IRISPen Executive users. You could for instance scan a text and "proofread" it at the same time.

With the IRISPen Translator, speech synthesis has a special meaning: to learn how to **speak a language**, you can just scan a text and hear the PC pronounce the words!

Summing up, your IRISPen uses a truly human way of interfacing. After all, isn't speech the most natural way of interacting with a machine?

As is to be expected of a multilingual system, up to 10 languages can be supported.

PREPARING FOR SPEECH SYNTHESIS

Click on the "Speech" button on the toolbar (or enable the option "Speech" under the "Options" menu) to activate the speech synthesis.

Doing so disables the option "Sound" automatically: it doesn't make any sense to mix real speech and system sounds!

Evidently, your PC must be equipped with a Windows compatible **sound card** if the speech synthesis is to work. Do not enable the speech feature unless you really intend to use it: the recognition slows down as you wait for the recognized data to be read aloud.

Enable the **multiline reading** mode to ensure a natural prosody. (Multiline reading is disabled by default.) When the text-to-speech module has to pronounce individual lines - "chopped up" text bits - rather than full sentences, the intonation may be somewhat less than "human".
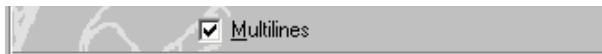
Unless additional language modules were installed, the speech library of the IRISPen Executive is limited to the language selected by the user at installation time. The IRISPen Translator is equipped with the speech libraries of the two languages in the language pair. For instance, the English-French version of the IRISPen Translator has both the English and French speech libraries. You are warned when you try to read a different language.
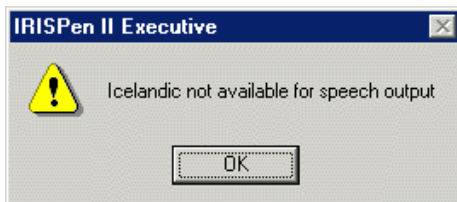
Users can extend the speech capability by adding additional speech modules. Ten languages are currently available: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.
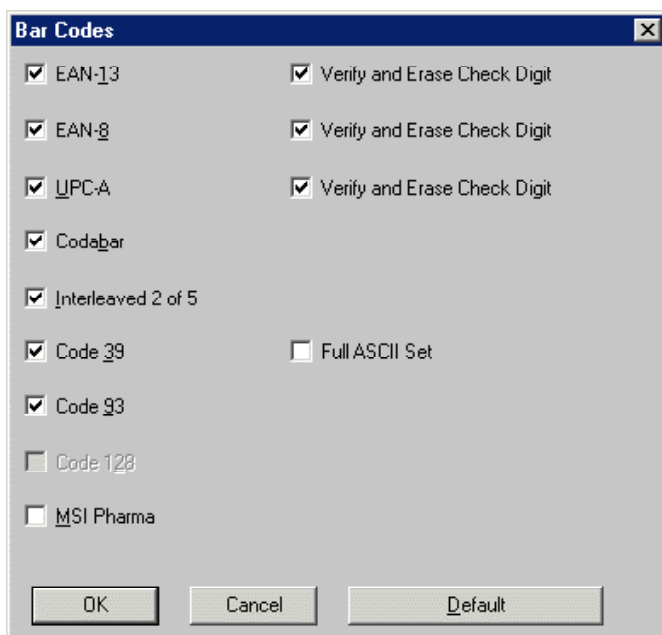
REPEATING THE SPEECH

You can repeat the last utterance at any moment: right-click the button "Speech" in the IRISPen toolbar (or use the command "Repeat Speech" under the "Options" menu). You can also define a repetition hotkey with the command "Hotkeys" under the "Settings" menu.
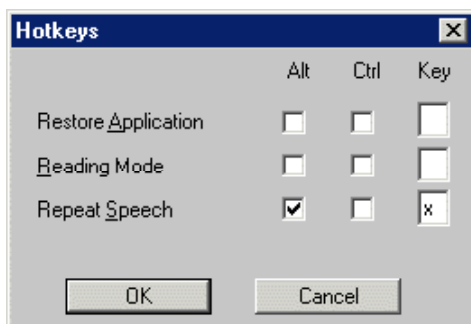
A WORD ABOUT TEXT TO SPEECH

On the whole, the operation of the text-to-speech feature is very easy and straightforward. But does this mean that the technology is as well? In fact, the opposite is true! Let's have a closer look at this exciting technology.

Converting text to intelligible sound is a complex process incorporating three steps: a linguistic, a phonetic and an acoustic phase. Each module plays a distinct roll in the text-to-speech process.

From Writing to Sounds

In the first stage, the linguistic module analyzes the recognized text.

Text normalisation is the first step. Punctuation symbols - comma's, dots, exclamation and question marks, hyphens etc. - are identified and abbreviations, dates, telephone numbers, account numbers etc. are converted into plain text.

Depending on the context, the punctuation symbols can get pronounced: "/" will be pronounced as "slash", ":" as "colon" etc. This only occurs when these symbols have to be pronounced to get noticed. Pronouncing special symbols is necessary to read structured character strings as "1245AS**/**876" but not in body text such as "Shakespeare**'**s Globe" and "John said**: '**I won**'**t.**'**" A typical example is reading telephone numbers such as "(800) USA**-**1234" where the "-" symbol is pronounced as "dash".

It is interesting to note that in some languages - for instance English, French and Dutch - the pronunciation of figures varies when you are dealing with a year and not a "normal" figure! The *number* 1,999 becomes "one thousand nine hundred and ninety", the *year* 1999 is pronounced as "nineteen ninety six".

Next, all **idiosyncrasies** of the recognized text are identified to determine the correct pronunciation. This means that the word "Dr." will be correctly interpreted as "doctor" and "drive" in sentences as "*Dr.* Spencer visited a patient on Roosevelt *Dr."*.

All common **abbreviations** are solved with a specialized dictionary: the system knows that "Corp." stands for "Corporation", "Bros." for "Brothers" etc. Lin-

guistic rules are applied to all other abbreviations, depending on their structure: "NASA" is pronounced as any "normal" word but abbreviations such as "Ph. D." and "MBA" will be pronounced letter per letter.

Abbreviations composed of the first letter of each word and pronounced as normal words are called *acronyms*: IRIS, short for "Image Recognition Integrated Systems" is an example. *Initialisms* such as "PC" and "VAT" are also composed of the first letter of each word but are pronounced letter by letter. To ensure that a word is read letter per letter, it suffices to insert dots between the letters: "N.A.T.O." for example yields a different result than "NATO".

Next, the conversion of **spelling to phonetics** takes place. Using a rule-based method, the letters are replaced by sounds, words are cut up in syllables and the necessary accents are identified. In short, the ortography is replaced by an accurate phonetic transcription.

This is by no means an easy task, as the English language perfectly illustrates. Some letters, such as x, c and q do not correspond to sounds! In the French language, the /o/ sound can be spelled in ten different ways: au, aux, eau (bureau), eaux (bureaux), os (dos), aut (haut), ault (Renault), o (zéro), ot (abricot) and ôt (plutôt). The playwright G. B. Shaw once jokingly proposed to write the word "fish" as "ghoti": "gh" is pronounced as /f/ in "enough", the "o" is pronounced as /i/ in "women" and "ti" is pronounced as /sh/ in "nation".

As individual sounds, so-called "phonemes", and not words, are the basic segments of a phrase, the introduction of new words and technical terms unknown to the system poses no threat. Shakespearean monologue and brand new words are pronounced with equal ease.

To solve **ambiguous cases** such as the *noun* "*re*cord" and the *verb* "re*cord*" in phrases such as "Let's re*cord* a beautiful *re*cord" or the *verb* "live" (/liv/) and the *adjective* "live" (/laiv/) in "The host of the live show lives in New York", a complete linguistic analysis of the recognized text on sentence *and* word level is performed. (The technical term for such words is "homographs": they are written identically, but pronounced differently.)

This **syntactic** and **lexical analysis** makes extensive use of general rules and dictionaries containing exceptions. It takes contextual linguistic knowledge to detect that "re*cord*" is a verb and "*re*cord" a noun. Dictionaries are again used to solve such cases: suffixes such as "-ness", "-logy", "-ly" can help to determine the word status.

From Isolated Sounds to Real Speech

The linguistic analysis and conversion of written symbols into sound equivalents is by no means sufficient: these segments have to be concatenated successfully to form fluent, coherent speech. Who pronounces isolated sounds one after the other with equal stress and volume for each one of them? You wouldn't even understand what was being said!

The **phonetic** module calculates precise speech patterns from the phoneme output of the linguistic module. To assure the smooth, non "robotic" character of the produced speech, the phonemes are replaced by sound segments originating with **native speakers**. Although you are dealing with a "synthesized" voice, the characteristics of the speaker's original voice are still recognizable in the spoken text.

This does not mean that individual phonemes are replaced in each case; **sound clusters** are also used for a more natural coarticulation, for better transitions between the individual phonemes. "Diphones" consisting of two sounds as /ai/ in the word "hide" and "triphones" consisting of three sounds as /aie/ in "buyer" are extensively used.

Other elements which must be added with native speaker-based models are a natural **prosody** - *intonation* and insertion of *pauses* - of each phrase and the **length** and **pitch** of each sound. (And that's why we recommend you to enable the multiline reading mode!) Here again, native speaker-based models are used for optimal results.

In most languages, at least one important word in a sentence gets an **accent**. Think of sentences such as "*John* went home" - where we mean it wasn't Peter - and "John went *home*" - and not elsewhere!

The *length* of a sound depends on many factors: the intrinsic length of the phoneme - the length of the /oo/ sound in "book" and "groovy" is very different - , the stress - stressed syllables are higher in *pitch* and longer - and some other factors.

Important *syntactic borders* are marked by a higher intonation, sounds are longer just before a syntactic border. The *type of sentence* also plays a role: it makes a substantial difference whether a sentence ends on a question mark, an exclamation mark or point as in declarative sentences.

Additionally, the characteristics of spoken words are identified, providing separate pronunciations for "rough", "cough", "dough", "plough", "though", "thought", "through" and "thorough".

Turning Speech Patterns into Sound Waves

The last phase of **acoustic processing** takes care of the actual speech production: the processed speech patterns from the phonetic phase are converted into synthetic waves which are reproduced by the PC's audio card.

EXPERIMENT TO LEARN!

To get a better idea of the highly advanced nature of the speech feature, an I.R.I.S. test sheet is provided with the IRISPen.

It allows you to experiment with the text-to-speech capability of your IRISPen Executive and get aware of the linguistic rules you should respect when you want to correctly read telephone numbers, bank account numbers etc.

LEARNING MORE

These **syntax rules** are discussed at length in Appendix B. They describe how the speech synthesis deals with normalized character strings such as bank account and telephone numbers, date and time indications etc.

To ensure that the speech module will produce correct results for such character strings, you should respect these rules when your documents - forms, labels etc. - are designed and produced. In other words, these **syntax rules** tell you how English dates, French telephone numbers and German bank account numbers should be composed to be recognized as such.

Chapter 4
BAR CODE READING

Bar code reading adds numerous applications for IRISPen owners: information desks, marketing surveys, customer identification, order picking and many other examples of data entry in the office environment can benefit highly from a bar code reading solution.

Where extreme accuracy and high speed are mandatory in identification processes, bar codes undoubtedly offer a great solution. The IRISPen Executive bar code reading is a powerful but very user-friendly data entry solution offering seamless integration of bar code reading into any Windows application. Identifying customers and products with laserprinted or inkjet printed bar code labels becomes a cinch.

As bands of about 0.4" (1 cm.) high are scanned, the IRISPen is more flexible than other bar code reading pens. You do not have to position the IRISPen with great precision to read bar codes correctly; skewing and image distortion are dealt with beautifully. Also note that the IRISPen is a greyscale pen scanner that allows you to adjust the scanning brightness, a unique feat for bar code scanners! As a result, a wider range of bar codes can be read, including bar codes on a colored background.

Although all common symbologies are supported, its use is transparent: no programming or parameter setting is necessary.

HOW DOES THE IRISPEN WORK ON BAR CODES?

This much we already know: the pen-sized scanner functions as a greyscale camera that sends an image to the computer. To recapture printed or typed infor-

mation, you simply glide the pen over a line of text. As you scan, an image consisting of intense points (so-called "pixels") on a lighter background is generated and transmitted to the PC. The IRISPen software now takes over.

All this applies to bar codes as well. The one difference is in the processing: the image captured with the pen scanner contains black and white bars, not characters, and the software analyzes these bars, not characters shapes, to extract a sequence of symbols.

Bar codes are composed of parallel bars and spaces between them. Predefined combinations of bars and spaces represent specific characters. The structure of a bar code is actually very similar to that of a Morse coded message: where Morse code represents characters with dots and dashes, bar codes represent symbols with successive dark and white bars.

There are several bar code standards such as Codabar, Interleaved 2 of 5, Code 39 etc. Appendix C of this manual discusses these "symbologies" in greater detail and contains some samples of bar codes the IRISPen can read with great ease.

Summing up, the IRISPen software analyzes the succession of dark and white bars to detect the characters the bar code contains. The PBR technology from I.R.I.S. - PBR is short for "Pen Bar Code Reading" - analyzes low-quality images which are the result of manual, unstable scanning; skewed, wavy and distorted scans are dealt with beautifully.

The sequence of extracted characters is, as when you read text information, inserted at the cursor position of the active Windows application. Use the IRISPen's bar code reading and you can enter bar code data into all Windows applications without any integration effort.

PREPARING FOR BAR CODE READING

To read *bar codes*, and not text, it suffices to select the **reading mode** "Bar Codes" in the IRISPen application window. Evidently, no language selection or other text-oriented settings apply to bar codes.
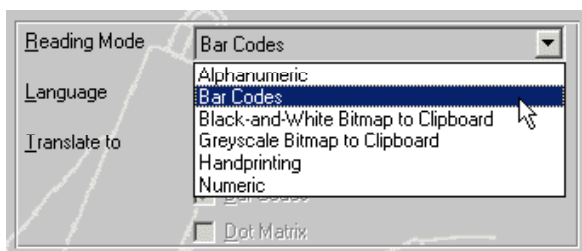
The selection of the reading mode is confirmed by the IRISPen status window.

That's all it takes, you are now ready to scan bar codes! Contrary to most bar code readers, you do *not* have to indicate which bar code **symbology** you are going to read. The IRISPen Executive software detects fully automatically if you are reading Code 39, Codabar or Interleaved 2 of 5 etc. type bar codes. As far as the user is concerned, the symbology is transparent; various bar code types can be mixed without any consequence.

FINE-TUNING THE OPERATION

Nevertheless, the IRISPen Executive software allows you to fine-tune the operation of the bar code reading. The user has full control over the used symbologies, should he desire it.

Click on the "Bar Codes" button on the toolbar to define the bar code settings.

Clicking on this button enables the reading mode "Bar Codes" at the same time (unless the mixed reading mode is enabled). If you want to activate this reading mode and adjust the bar code settings *simultaneously*, a click on this button suffices.

By default, all bar code symbologies are enabled and the IRISPen will autodetect the symbology, but you can limit the recognition to specific bar code types. Simply click on the corresponding options to enable and disable a specific symbology.

Bar Codes

☑ EAN-13	☑ Verify and Erase Check Digit
☑ EAN-8	☑ Verify and Erase Check Digit
☑ UPC-A	☑ Verify and Erase Check Digit
☑ Codabar	
☑ Interleaved 2 of 5	
☑ Code 39	☐ Full ASCII Set
☑ Code 93	
☐ Code 128	
☐ MSI Pharma	

OK   Cancel   Default
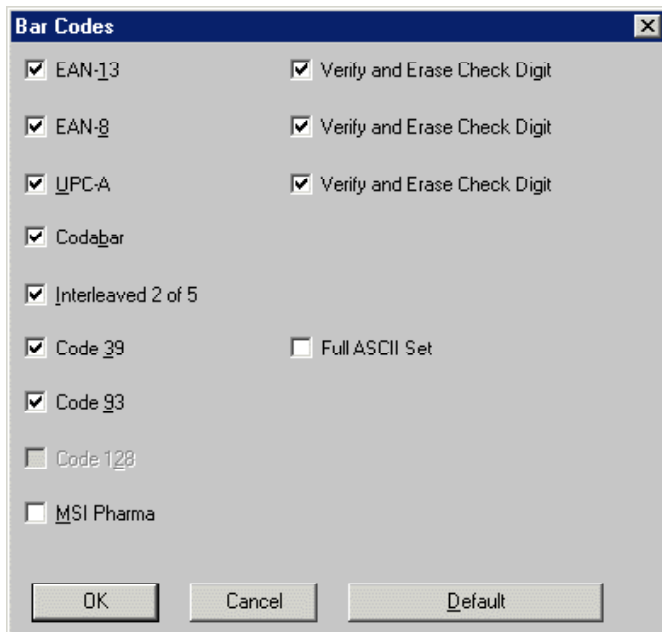
For the symbologies EAN-13, EAN-8 and UPC-A, you can choose to verify the **check character(s)** with the option "Verify and Erase Check Digit". When this option is enabled, the check digit is verified but filtered from the output. When this option is disabled, the check character is not verified but included in the output. Disable this option to support the variation of the UPC-A symbology which does *not* contain a check digit.

Another option concerns the Code 39 symbology. An extended version of Code 39, a symbology which in its standard version has a limited character set, allows to encode all ASCII characters. The option "Full ASCII Set" allows to interpret that symbology. By "full" ASCII set, we mean all characters up to character 128 in the ASCII table. In simple words, the "special" and accentuated symbols from 128 onwards such as ç and ó are *not* included.

The "Default" button restores the default settings. It should be noted that the bar code settings are included in the IRISPen settings! The command "Save Settings" under the "File" menu and terminating the software with the option "Save Settings on Exit" under the "Options" menu enabled include the selected values in the IRISPen settings.

READING BAR CODES ALONGSIDE OTHER DATA

You can read bar codes and other data *at the same time*! You do not have to switch from one reading mode to another to encode documents which contain bar codes and another type of information.

The parameter "Reading Mode" in the IRISPen application window determines the primary reading mode, the option "Bar Codes" in the IRISPen application window determines whether you read bar codes *at the same time*. Enable this option to use the mixed reading mode.

Reading Mode	Alphanumeric	▼
Language	English	▼
Translate to	<None>	▼
	☑ Bar Codes	
	☐ Dot Matrix	
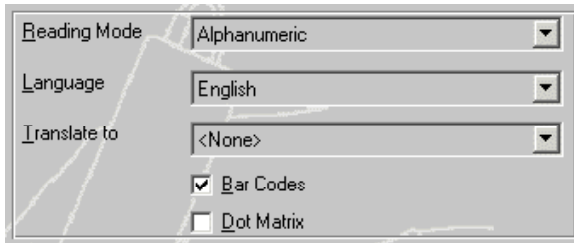
The IRISPen will "auto-detect" the type of data that is being read. You can for instance scan address labels where the printed data is complemented by a bar code without ever changing the reading mode!

SCANNING TIPS

Let's try and read some bar codes now! Use the I.R.I.S. bar code test sheet to this end. Select the correct IRISPen settings, click the "OK" or "Minimize" button to make the IRISPen application run in the background and start up any text-based Windows application.

Try to maintain a straight line and constant speed during your scan. The **line guide** on the view scanning window should be placed in the middle of the bar code. The more regularly you scan, the better your chances are of recognizing the bar code.

The greyscale image window shows the greyscale scan, the scan test window shows the binarized image and executes the bar code reading. Use these image windows to judge the quality of the images you produce and improve your scanning skills by trying different angles and speeds of scanning.
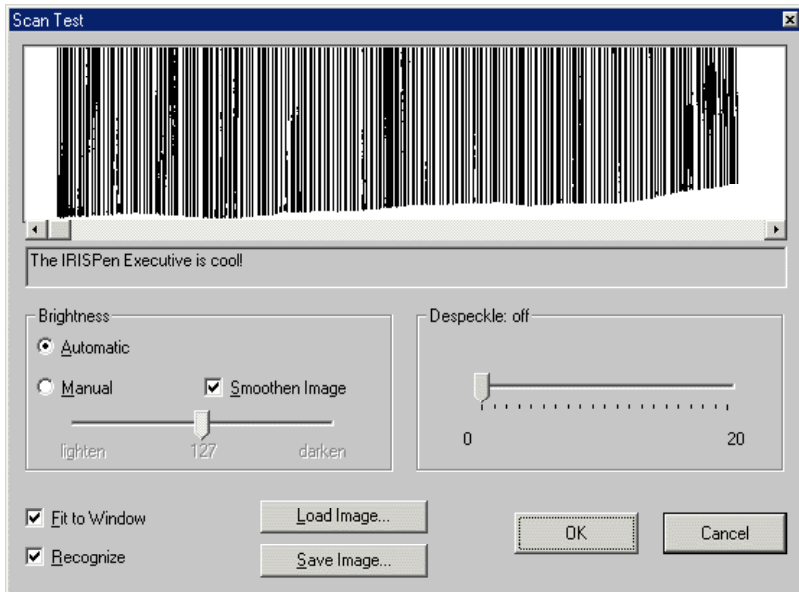
Always check that you start scanning clearly *ahead* of the bar code and end well *after* it; most bar code types require a "quiet zone" in front of and behind the actual bar code. Partial scans will not render partial results!

All bar code generating software packages provide an option where a large "**quiet zone**" is added to the actual bar codes. Enable this option, or use the IRISPen's "drop out" colors red and yellow as environment, as background color for the bar codes - these colors won't be picked up by the pen scanner anyway! (This also implies that all information printed in red and yellow is not legible.)

It is interesting to note that some symbologies, UPC-A and EAN-13, are designed in such a way that it makes no difference whether you scan them from left to right or right to left!

Never read more than a **single bar code per scan**! Some product labels are composed of two bar codes, where the first bar codes identifies the product and the "supplemental" bar code specifies the retail price. When a single image contains two bar codes, the widest bar code gets read and the secondary is ignored, as the example below illustrates.

GOOD READING OF BAD IMAGES

You will soon experience, that, as much as you might want to, it is very difficult to scan with a perfectly steady hand. So, the obvious question arises whether this undermines the accuracy of the bar code reading.

It doesn't! Evidently, bar code reading with the IRISPen has its limits - these are discussed below - and one cannot expect the same flexibility from the IRISPen that is typical of bar code laser scanners. The IRISPen is a bar code reader for the office environment, not for industrial applications.

Nevertheless, it is important to stress how tolerant the IRISPen is with images of bad quality. As bands of about 0.4" (1 cm.) are scanned, the IRISPen is more flexible than other bar code reading pens. Even when your scanning generates skewed, wavy and distorted images, the proprietary PBR technology is able to correctly read the bar codes! You do not have to position the pen scanner with great precision to read bar codes correctly.

Below are some typical examples of images you can generate when you scan less than perfectly.

While the first bar code is read at a constant angle, thus producing a skewed image, the second bar code has become wavy because of a varying scanning direction.
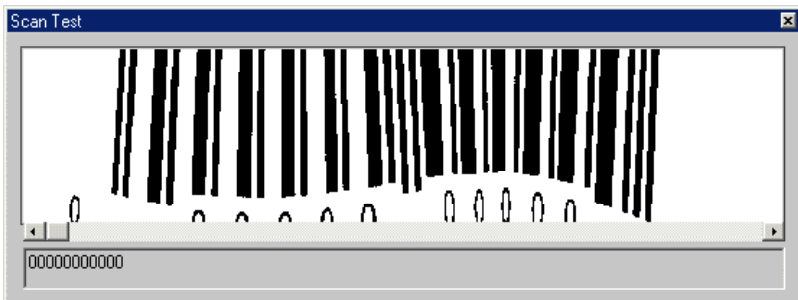
SETTING THE BRIGHTNESS

The test mode is also useful to set the brightness correctly. Unlike other bar code scanners, the IRISPen is a greyscale scanner that allows you to adjusts the scanning brightness to the color tones of your forms, labels etc. As a result, a wider range of bar codes is legible!
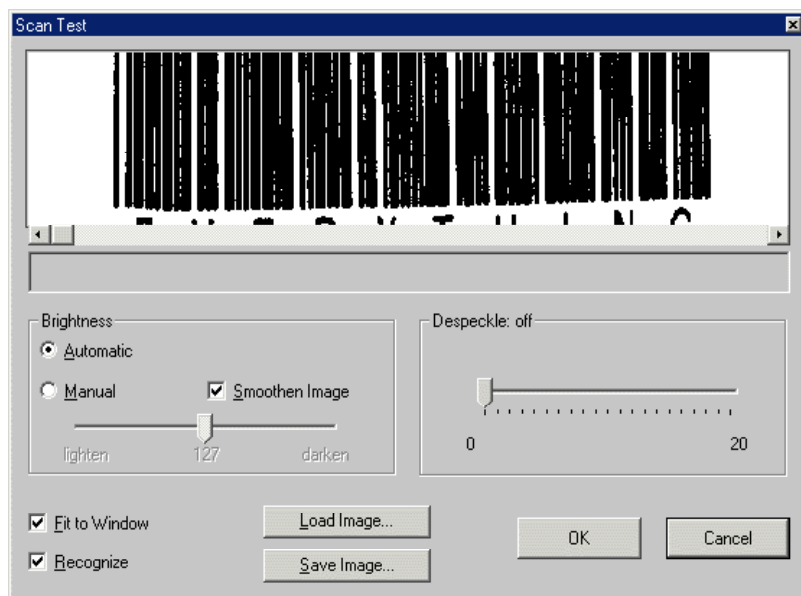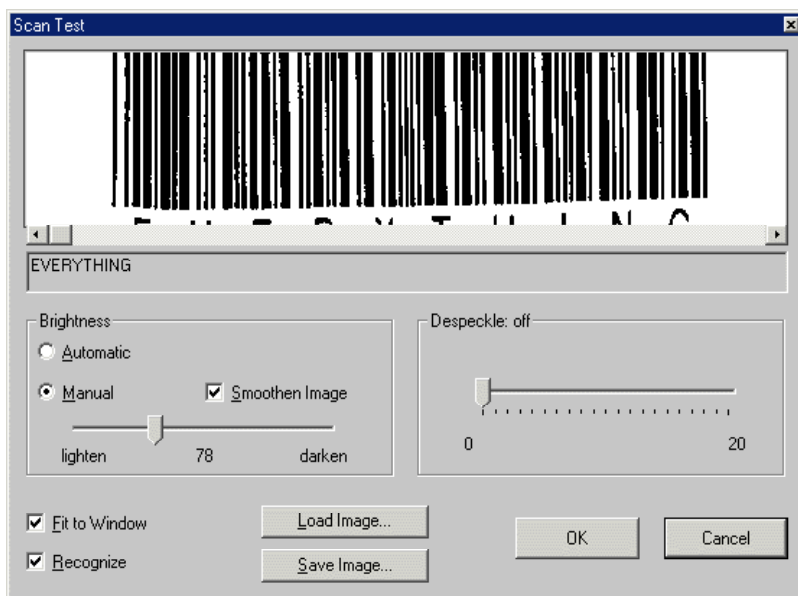
To provide optimal security, bar codes never yield a partial result: a skipped bar, a missing start or stop character and a incorrect check digit always lead to a misread, a zero result.

ADDING AUDITIVE FEEDBACK

It goes without saying that you can activate the speech feature to get auditive feedback of your bar codes. Speech synthesis not only applies to text but to all legible data: bar codes are after all nothing but characters encoded as black and white bars! For the "hands- and eyes-busy" activity of reading bar codes, auditive feedback is certainly an option.

Refer to Chapter 3 for more information.

WHICH BAR CODES CAN THE IRISPEN READ?

Let's discuss the IRISPen's bar code reading capability in depth now that we are familiar with its operation. Which bar code symbologies can the IRISPen software read? What limitations does the IRISPen scanner have?

As far as **color tones** are concerned, the IRISPen works best on well contrasted bar codes. Black bars on a white background yield the best results. (We repeat that you can use the image adjustment features to adapt the scanning to your documents and that red and yellow are the IRISPen's invisible or "drop out" colors.)

Secondly, the **printing quality** is relevant: only laserprinters and inkjet printers provide sufficient quality. Matrix printers should be excluded as they do not produce sufficient sharpness - their resolution is mostly limited to 60 dpi - and contrast - even when the brightness is adjusted. Here's an illustration: this bar code gives you the following black-and-white image when scanned.
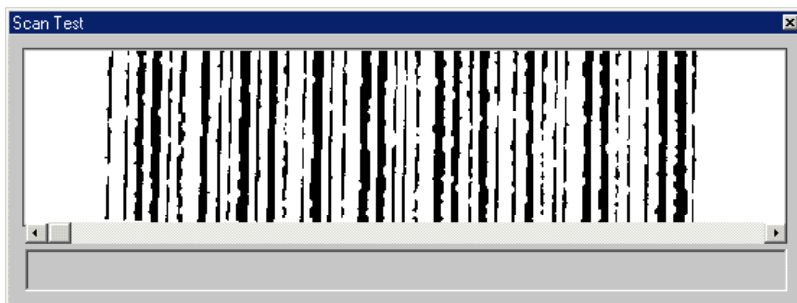
Furthermore, you are recommended to scan on a flat, stable **surface**. Although you will undoubtedly succeed in reading some bar codes on physical objects such as bottles, boxes etc., these are not part of the intended IRISPen bar

code applications. The IRISPen is a productivity tool for the **office environment**, not a scanner for industrial applications! It would therefore be unrealistic to expect the same flexibility from the IRISPen one takes for granted for a bar code wand or laser scanner.

The IRISPen supports **bearer bars**. "Bearer bars" are lines drawn around a bar code which have two functions: they enhance the printing quality of bar codes by guaranteeing that the bar code (label) printer exerts equal pressure on the entire surface of the bar code.

They also enforce some scanning discipline: when the scanning is too skewed, you cross a bearer bar as you enter or exit the bar code and thus provoke an invalid result.

The **height** and **width** of the bar codes also play a role. For a bar code to be legible by the IRISPen, the following conditions should be met: the bar code should be at least 0.3" (0.75 cm.) high and the narrow bar should be at least 12 mils (0.012" or 0.3 mm.). When tinier bar codes are used, the **resolution** of the pen scanner does not suffice.

LEARNING MORE...

Appendix C of this manual discusses the "symbologies" or bar code standards the IRISPen Executive software can read and contains some samples of bar codes the IRISPen can read with great ease.

Chapter 5
HANDPRINTING
RECOGNITION

Last but not least, the IRISPen Executive also offers handprinting recognition. Financial applications, form reading, encoding labels etc. can benefit greatly not only from bar code reading but also from recognition of handprinted numerals.

This proprietary recognition technology from I.R.I.S. is based on more than 1 million writing samples! As a result, all natural writing styles - American and European - are supported. Contrary to many handwriting recognition solutions, no training phase or imposed style is required.

User control over spacing and rejection levels is added for flexibility.

PREPARING FOR HANDPRINTING RECOGNITION

To read *handprinted* numerals, and not printed or typed text, it suffices to select the **reading mode** "Handprinting" in the IRISPen application window.
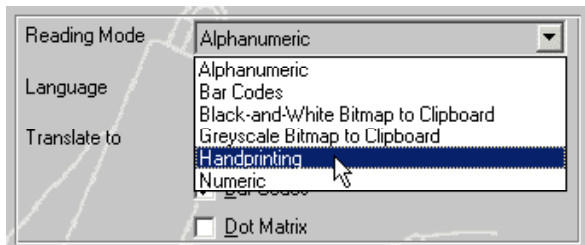
The selection of the reading mode is confirmed by the status window.

That's all it takes, you are now ready to scan handprinted numerals!

HANDPRINTING OR HANDWRITING?

Up to now, we used the word "handprinting" to describe the recognition of handwritten symbols by the IRISPen. What do we mean exactly by hand*printing*?

Handwriting is used to describe continuous or "cursive" handwritten text. The symbols within a word or character string are linked, it is impossible to say where one symbol ends and another starts. With **handprinting** on the other hand, the symbols are separated and the recognition software has an easier job to isolate the individual characters.
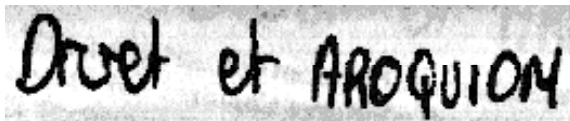
To strengthen the recognition results further, the IRISPen limits recognition to the **numeric symbols**: the numeral '8' cannot be mistaken for a "B", '0' cannot be confused with an "0" or "o" etc.
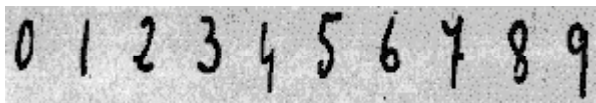
The dot, comma and hyphen symbols are also recognized for **banking applications**. However, spaces are ignored in the output. The characters are "glued", no word concept is used when handprinted symbols are read. On the other hand, telephone numbers, bank account numbers, dates etc. are recognized by the speech module. When the speech feature is enabled, the word "area code" is added in telephone numbers, pauses are inserted where dashes occur in bank account numbers etc.

FINE-TUNING THE OPERATION

The IRISPen Executive software allows you to fine-tune the operation of handprinting recognition. Click on the button "Handprinting" on the toolbar to modify these settings.

Clicking on this button enables the reading mode "Handprinting" at the same time. If you want to activate this reading mode and adjust the handprinting settings *simultaneously*, a click on this button suffices.

You have full control over segmentation and the rejection level.
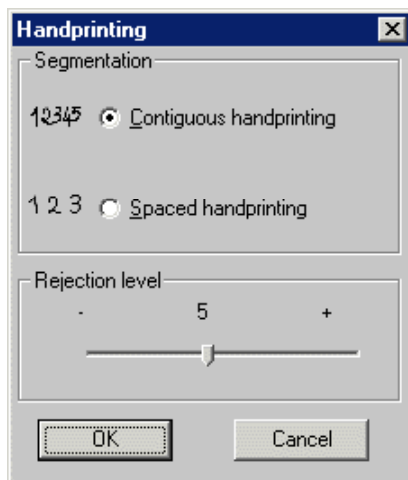
First, you can tell the system if the characters are well spaced or contiguous. Two examples make the difference clear. The example below shows a spaced character string.
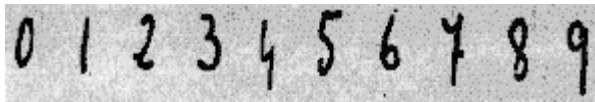
The second example illustrates contiguous spacing. The space between two characters is considerably reduced, some characters may occassionally touch each other and the zones covered by a single character may overlap. Nevertheless, the IRISPen recognizes the bank account number perfectly.
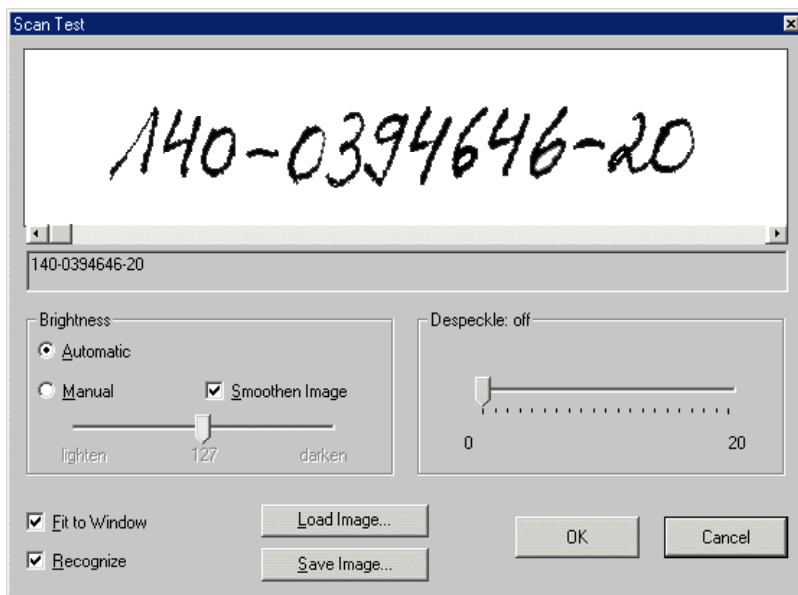
It is mandatory to set this parameter correctly. When you try to read contiguous handprinting with the segmentation set to spaced, some characters may be skipped by the system!

The easiest way of enforcing spaced printing is by developing **forms** where the individual numerals must be placed in predefined blocks. To ensure that these

predefines zones are not picked up by the IRISPen, the invisible or "drop out" colors red and yellow should be used.

Secondly, the user has influence over the rejection rate. Drag the slider bar to adjust the rejection level. When the rejection threshold is increased, characters will be marked as unknown symbols quicker, there will be less substitutions; when the rejection rate is decreased, the system will substitute more symbols as it tries harder to find a match.

For applications where extreme accuracy is essential, you should maintain a high rejection level. Evidently, when you are encoding structured data such as client numbers etc., it is always useful to integrate a check digit which is verified by the encoding utility.

A WORD ABOUT THE TECHNOLOGY

You cannot read handprinted numerals and other data *at the same time*. The nature of the images to analyze is very different, as the graphics below illustrate.

The first time, a typed bank account number was scanned. Although an old typewriter was used, the individual symbols are nicely separated and each font shape is clearly legible as all bars, loops are sufficiently thick etc.
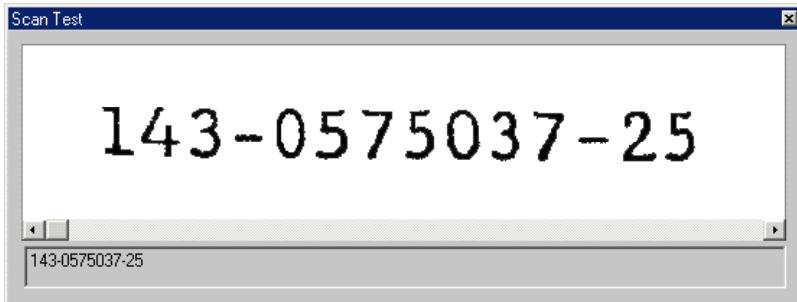
The second time, a handwritten character string was scanned. All characters are linked as human beings write in a continuous pen movement. Even for the block letters on the right - these *are* separated - the image is of much lower quality than is the case with printed and typed material. As you write, the pressure of your hand changes constantly; inevitably, some parts are written much lighter than others. Moreover, in most cases pens are used which do not provide the necessary thickness.
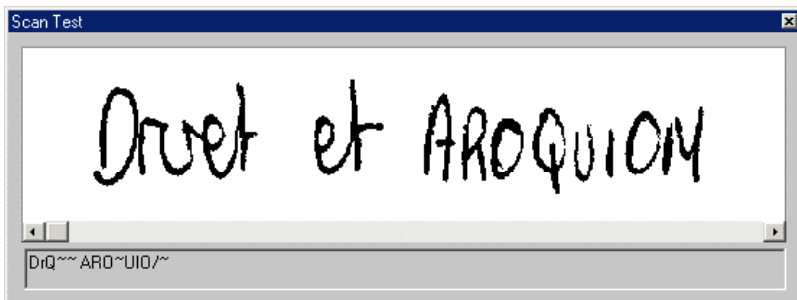
Unrecognized symbols are by default represented by a tilde (the "~" symbol).

Given the dramatically different nature of the images to be analyzed, it is easy to understand that it takes highly specialized recognition software to recognize handwritten and handprinted symbols!

The proprietary handprinting technology from I.R.I.S. is based on more than 1 million writing samples. All natural writing styles are supported: American and European ones and sevens are recognized with equal comfort. It requires no training phase or imposed style to recognize somebody's handwriting.

M*IXING* H*ANDPRINTING AND* O*THER* D*ATA*

A direct consequence is that you have to switch from handprinting to the other reading modes to encode documents where more than one type of information is being read.

(However, you can read handprinted data and bar codes and other data *at the same time*! The parameter "Reading Mode" in the IRISPen application window determines the primary reading mode, the option "Bar Codes" in the IRISPen application window determines whether you read bar codes *at the same time*.)

Switching from one mode to another is very easy. You can obviously do it by modifying the reading mode in the IRISPen application window. There are several ways of accessing the IRISPen application window quickly: right- or double-clicking the IRISPen status window or right-clicking the IRISPen control on the Windows taskbar and selecting the command "Open". You can restore the IRISPen application directly from within your target application by defining a "restore" hotkey with the command "Hotkeys" under the "Settings" menu.

The same command allows to define a **hotkey** to switch from one reading mode to another directly from within your target application.
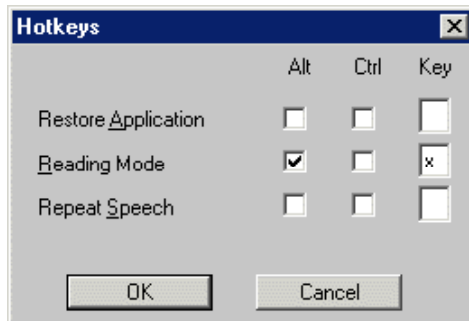
Let's give an example: we define Alt+X as the hotkey for reading mode switching. For starters, we select the setting alphanumeric as our reading mode. We now go to our Windows application and scan in some text. When we execute the keyboard command Alt+X inside our Windows application, the IRISPen switches to bar code reading. When we repeat the keyboard command, the reading mode successively changes to black-and-white and greyscale bitmaps to clipboard, handprinting, numeric and back to alphanumeric. The reading mode is always specified in the status window!

Secondly, we can use **pen button** clicks to change the reading mode. To encode documents where handprinting and other information are mixed on the page, just clicking the pen button can do the job. Click on the "Pen Button" button on the toolbar.

A single and double pen button clicked can be programmed to switch from one reading mode to another.

Pen Button

Single		Double
○	Null	○
○	<Enter>	○
○	<Space>	◉
○	<Tab>	○
○	<Cursor Up>	○
○	<Cursor Down>	○
○	<Cursor Left>	○
○	<Cursor Right>	○
○	<Home><Cursor Down>	○
○	Custom...	○
◉	Switch reading mode	○

OK   Cancel
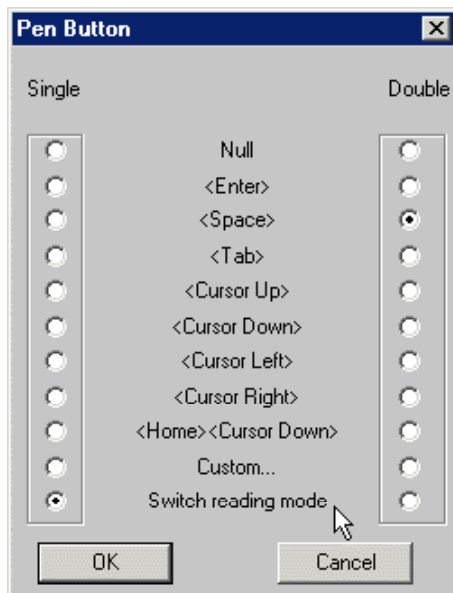
In our example, the single pen button click will execute the switching of the reading mode.

ADDING AUDITIVE FEEDBACK

It goes without saying that you can activate the speech feature to get auditive feedback of your handprinted numerics. For the "hands- and eyes-busy" activity of reading handprinting, auditive feedback is certainly an option.

It should also be noted that the speech feature incorporates intelligent routines that allow it to recognize bank account and social security numbers, telephone numbers, dates etc. When you are reading bank account numbers, short pauses are inserted where a dash occurs, when you are scanning telephone numbers, the

speech synthesis will add the word "area code" where it is useful etc. These routines are language-dependent; no less than 10 languages can be supported!

Refer to Chapter 3 for more information.

LEARNING MORE

Where less than optimal reading results are obtained, it may be useful to adapt your writing style and optimize your forms to maximize your productivity. Appendix D discusses a few simple rules that can teach you why substitutions occur and how to avoid them.

Chapter 6

COMMAND REFERENCE

This chapter provides an extensive reference guide to the commands of the IRISPen, the IRISPen Executive and the IRISPen Translator. Also note that buttons are available for all frequently used functions!

IRISPEN APPLICATION WINDOW

The IRISPen application window contains the basic IRISPen settings: the reading mode, the document language, the target language for translation, the enabling of bar code reading as secondary reading mode, the target output and the multiline reading mode.
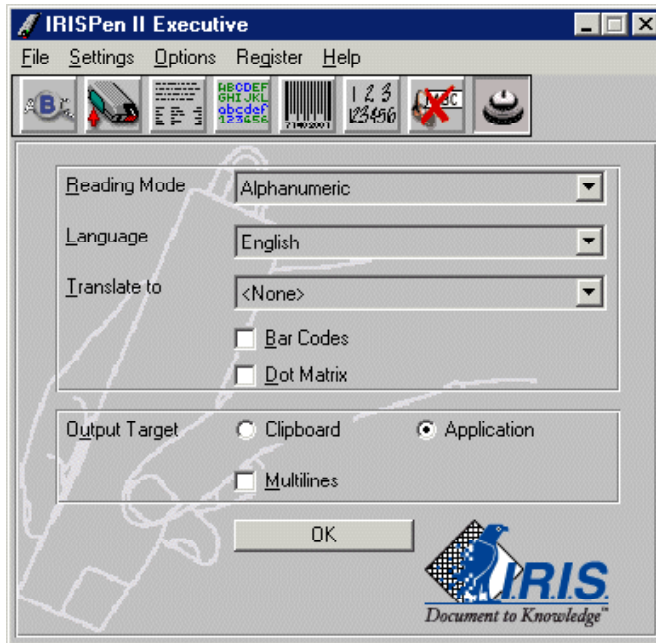
The selected values are displayed in the IRISPen status window. Whether the status window is displayed or not depends on the option "Status Window" under the "Options" menu.

The "OK" or "Minimize" button is clicked to have the IRISPen application running in the background.

To restore the IRISPen application, the user can right- or double-click on the status window, and he can right-click on the IRISPen control on the Windows taskbar and select the command "Open". He can also define a shortcut key to this end with the command "Hotkeys" under the "Settings" menu. Such a shortcut

restores the IRISPen application window from within the user's active application.

Reading Mode

Determines the reading mode.

The options alphanumeric, black-and-white bitmap to clipboard, greyscale bitmap to clipboard and numeric are available. For the IRISPen Executive, two reading modes are added: bar codes and handprinting.

For translation help and language learning with the IRISPen Translator, the user should obviously select the option "Alphanumeric".

For financial applications, the set of recognized characters can be limited to numeric symbols, in which case only the characters '0' to '9' and some punctuation symbols such as '.', ',' etc. will be recognized. Any chance of recognition errors such as the confusion between '5' and "S", '0' and "O", '9' and "g", '8' and "B" etc. is eliminated.

Fully customized character sets can be programmed by the user with the "Output Subset" button on the toolbar. An example is creating an OCR-B character set which limits recognition to the numerals and the symbols "<", ">" and "+". (OCR-B is at the same time a highly specialized typeface used only in financial documents such as preprinted checks, bank statements etc. I.R.I.S. has a specialized version of the IRISPen for the financial market.)

Selecting a limited character set not only increases recognition accuracy but also speeds up the recognition process somewhat.

The reading modes "Black-and-White Bitmap to Clipboard" and "Greyscale Bitmap to Clipboard" offer a direct way of inserting graphics in documents. Using the Windows clipboard, the user pastes the scanned image directly in his active application. After the scanning, a blinking status window displays the message "Image in clipboard" for a few seconds. As the image is sent directly to the Windows clipboard, it cannot be retouched.

(The user can also enable the scan test mode and save graphics in a Windows bitmap (BMP) file with the button "Save Image". In this case, he produces a graphic file which can be retouched and should be inserted in his documents later on.)

Selecting the reading mode "Bar Codes" or "Handprinting" suffices to scan bar codes and handprinted character strings. Contrary to most bar code readers, the user does not have to indicate which bar code "symbology" he is going to read. The IRISPen software detects fully automatically which bar code type has been scanned. As far as the user is concerned, the symbology is transparent and various bar code types can be mixed without any consequence. The handprinting technology supports all natural writing styles - American and European. Contrary to many handwriting recognition solutions, no training phase or imposed style is required.

There is an alternative way of selecting the bar code and handprinting reading modes. Unless bar code reading is already enabled as secondary reading mode with the option "Bar Codes" in the IRISPen application window, clicking on the "Bar Codes" or "Handprinting" button on the toolbar defines settings and selects the corresponding reading mode at the same time.

When bar codes and handprinted symbols are read, the language selection in the application window does not apply. Nor do the text-related commands - detection of table frames, character subsets and conversions etc.

The reading modes of the IRISPen Executive can be complemented by bar code reading as secondary reading mode with the option "Bar Codes" in the IRISPen application window.

The user can define a shortcut key with the command "Hotkeys" under the "Settings" menu and program the single or double click of the pen button with the "Pen Button" button on the toolbar to switch from one reading mode to another.

Language

Indicates the language of the text to be recognized.

Recognition by the IRISPen is not limited to English: 55 languages are supported! All American and European languages are supported, including the Eastern-European languages, Greek and Cyrillic. (The British and American - or should we say "international"? - variants of the English language are distinguished. The same goes for Spanish and Mexican.)
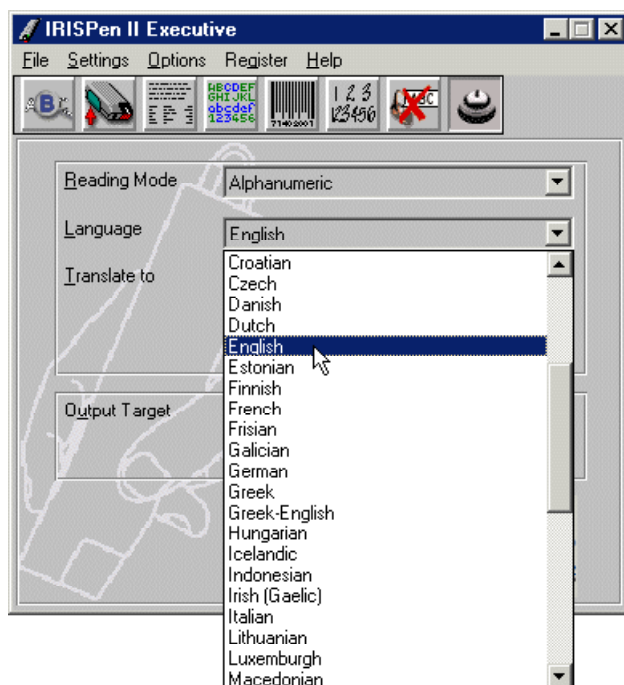
The IRISPen even copes with mixed alphabets of Greek or Cyrillic and Western characters. The software can switch languages in the middle of a sentence without any help from the user! When Western words pop up in Greek and Cyrillic - many untranscrible proper names, brand names etc. are written using the familiar Western symbols -, the IRISPen can switch to the correct alphabet automatically.

The user can press a letter key to move to it directly: if English is currently selected, and he wants to select Hungarian, he can click the "H" key on his keyboard to go directly to the Hungarian language. When several languages have

the same initial, he presses the letter several times to go through the options. Let's give an example: the IRISPen reads Polish and Portuguese. By pressing "P" once, the user selects Polish, by pressing "P" a second time, he selects Portuguese, and by pressing "P" a third time, he's back on Polish. (To go to *another* letter, say T, he presses BackSpace before he enters the "T" character.)

The status window indicates the selected language.

Selecting the proper language is imperative: based on the selection of a language, the software knows which symbols to recognize. Multi-linguistic support ensures that "special" characters such as ç, ß, ñ, γ and ø are recognized correctly.

Secondly, the software extensively uses linguistic data to validate its results. Suppose that the word "president" is scanned where an ink stain makes the "r" look like an "f". Looking things up in the English lexicon, the IRISPen software will detect autonomously that the word "president" is being read and that it doesn't make any sense to recognize the symbol "f". This "self-learning" technique is highly dependant on the linguistic context.

It takes the appropriate Windows configuration to display Central-European, Greek, Turkish, Cyrillic and Baltic characters. The user may have to install the Windows module "Multilanguage Support" before his Windows system is able to cope with these languages. The icon "Add/Remove Programs" under the "Control Panel" is selected to find out if this Windows module is installed on your PC.

To read exotic languages, the user also needs to enable the correct keyboard layout under Windows. By doing so, an appropriate font is activated. The Windows control panel "Keyboard" allows to install the necessary keyboard layouts.

To handle mixed alphabets, the user should select "Greek-English" or the appropriate Cyrillic language setting - for instance "Byelorussian-English". In other words: he should *not* select "Greek" or "Byelorussian" as document language and hope that the Western symbols will come out fine!

To mix other languages, the user selects the language with the most extended character set. If he has a document where the, say, French translation is placed

alongside an English text, he should select French as language to ensure that the accentuated characters such as ç, é and ù get recognized correctly.

For the IRISPen Translator, the language setting is used to indicate the original or "source" language. It takes two parameters to indicate the translation direction: the option "Language" indicates the source language, the option "Translate to" in the IRISPen application window indicates the target language. The user translates for instance English into French.

The language selection does not apply when bar codes and handprinted symbols are read and when scanned graphics are sent to the Windows clipboard.

The support of 55 languages does not imply that speech synthesis is available for all languages. (Up to 10 speech modules can be installed: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.)

Translate into (IRISPen Translator)

Indicates the "target" language of the text to be translated.

It takes two parameters to indicate the translation direction: the option "Language" in the IRISPen application window indicates the source language, the option "Translate to" indicates the target language. The user translates for instance English into French.

The status window lists the translation direction: when English is translated into French, the status window mentions "English>French".

The IRISPen Translator provides bidirectional, real-time sentence translation of all scanned text. The translation is displayed in a "pop-up" translation window. (The user can quickly click the "Copy" button of the translation window to make it disappear.)

The user clicks the "Copy" button of the translation window to transfer the translated text to other windows using the clipboard. He can paste the translation in the same window as the one where the scanned text is placed to study the foreign sentences and their meaning comfortably later on. He can also insert the

translation in a different window to save the original text and the translation in separate files.

Whereas this option enables the translation of scanned text, the option "Translate Clipboard" under the "Options" menu allows the user to translate text copied to the clipboard - text entered on the keyboard, wordprocessor files, mail messages received in a foreign language, Internet pages displayed with a web browser etc.

He can even send the scanned text directly to the clipboard - by selecting the clipboard as output target with the parameter "Output Target" in the IRISPen application window. The recognized text will not be inserted in the active application.

The translation feature is limited to the languages of the language pair. The IRISPen Translator English-French for instance translates English into French and vice versa. Extra translation libraries can be installed. The currently available versions are: English-French, English-German, English-Italian, English-Portuguese and English-Spanish. Contact I.R.I.S. to learn the availability of new language pairs.

The speech feature is available for all source and target languages.

The IRISPen Translator has sufficient linguistic intelligence to translate abbreviations, normalized character strings such as "1st", phrases consisting of several words such as "state of mind" (état d'esprit) and "take away" (wegnehmen) etc.

However, the user should be aware that computer-aided translation gives an imperfect first draft, an "information-only" translation. Should a translation be published, it will need to be refined by a human translator. The IRISPen Translator should be used to translate business-like, non-emotive language. As computers do not cope well with connotation, metaphor and irony, literature, poetry and some advertizing literature do not translate well, and often lead to humorous effects.

Related Commands

The option "Translate Clipboard" under the "Options" menu allows the user to translate texts copied to the clipboard - text entered on the keyboard, wordprocessor files, mail messages received in a foreign language, Internet pages displayed with a web browser etc.

Dot Matrix (IRISPen Executive)

Determines whether the user recognizes "normal" or dot matrix printed documents.

Setting this parameter correctly is mandatory: "draft" or "9 pin" dot matrix symbols are made up of isolated, separate dots, and highly specialized recognition routines are required to recognize them.

"Letter quality" dot matrix printing, also called "25 pin" or "NLQ" dot matrix, does not requires the enabling of this option. Nor do the printing qualities typeset, typewritten, laser printed and inkjet printed.

Bar Codes (IRISPen Executive)

Enables a mixed reading mode where bar codes are read alongside data of another type. The user can for instance scan bar codes and "normal" text without changing the reading mode.

The parameter "Reading Mode" in the IRISPen application window determines the primary reading mode, this option determines whether read bar codes are read at the same time.

Output Target

Indicates where the text result will be placed: recognized data can be inserted directly in the active application or transferred to the clipboard for later pasting. Normally, the recognized data gets sent to the active Windows application.

The clipboard can only contain the recognition result text of the last scan - with the understanding that several lines can be "buffered" when the multiline reading mode is enabled!

When the reading mode "Black-and-White Bitmap to Clipboard" or "Greyscale Bitmap to Clipboard" is enabled, no output target can be selected: the scanned images are always copied to the Windows clipboard for later pasting.

With the IRISPen Translator, the active application functions as the output target when the source text must be saved. The translation appears in the "pop-up" translation window, the "Copy" button of that translation window allows the user to paste the translation.

To study foreign sentences and their meaning comfortably later on, the translation can be inserted in the *same* window where the scanned text is placed; to save the original text and its translation separately, the translation can be inserted in a *different* window.

The clipboard functions as the output target when the user wants to concentrate on the translation without inserting the source text. Clipboard translation must be enabled with the command "Translate Clipboard" under the "Options" menu if this is to work. All scanned text is then sent directly to the clipboard, and translated automatically. The user can limit himself to inserting the translation in his application with the "Copy" button of the translation window.

Multilines (IRISPen Executive - IRISPen Translator)

Enables the multiline reading mode, where multiple lines are "buffered" before they are inserted, spoken and translated.

It takes a click on the pen button to "flush" the scanned text. After each scanned line, the status window displays the message "Press pen button to flush". The user should only do so at the end of a sentence or paragraph. He does not have to wait for this message to disappear from his screen to start scanning another line.

When this option is disabled, the output of each scan is transmitted automatically to the target application; the user does *not* have to press the pen button to start the output process.

Multiline scanning is recommended when the user reads text excerpts and when the speech synthesis is enabled. The multiline scanning mode ensures a

natural prosody. When the text-to-speech module has to pronounce individual lines, rather than full sentences, the intonation may be somewhat less than "human".

The multiline reading mode is simply mandatory for the IRISPen Translator, it is essential for language learning and translation aid. It takes a click of the pen button at the end of a sentence or paragraph to "flush" the scanned text, that is to insert it in the application (or copy it to the clipboard) and start the speech and translation process. Correct pronunciation and translation imply that full sentences are scanned, otherwise the system would pronounce partial sentences and (try to) translate ungrammatical text lines, with unavoidably weak results.

It should be noted that the pen button remains a programmable command button controlled by the "Pen Button" button on the toolbar. When multiline scanning is enabled, the first click of the pen button triggers the "flush", all consecutive pen button clicks are interpreted according to their programming (if any). An immediate *double*-click preceding the "flush" is also interpreted as the "flush" command.

When this option is disabled, the text output is transmitted automatically to the target application; the user does *not* have to press the pen button to start the output process.

Related Command

The basic settings can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

IRISPEN CONTROL ON WINDOWS TASKBAR

The IRISPen control on the Windows taskbar confirms that the IRISPen application is running.

It indicates whether the IRISPen is available - if the pen scanner is detached from your PC's USB port, a different icon is displayed.

The user can right-click the IRISPen control to call up several frequent commands: he can enable the greyscale image window, the status window and the scan test, restore and close the IRISPen application and get access to the on-line help.

<u>Open</u>

This command restores the IRISPen application which normally runs in the background. The user can now modify the various system settings.

<u>Related Command</u>

The user can define a shortcut key to restore the IRISPen application with the command "Hotkeys" under the "Settings" menu.

<u>Greyscale Image</u>

Enables the greyscale image window.

Refer to the section "Greyscale Image Window" for an extensive discussion of the gresycale image viewer.

<u>Related Commands</u>

The user can also enable the greyscale image window with the command "Greyscale Image" under the "Options" menu and by loading IRISPen settings with the command "Open Settings" under the "File" menu.

<u>Status Window</u>

Enables the status window.

Refer to the section "Status Window" for an extensive discussion of the status window.

Related Commands

The user can also enable the status window with the command "Status Window" under the "Options" menu and by loading IRISPen settings with the command "Open Settings" under the "File" menu.

Scan Test

Enables the scan test mode where the user displays a black-and-white version of the greyscale images generated by the pen scanner and the reading results.

Refer to the button "Scan Test" on the toolbar for an extensive discussion of the scan test mode.

Related Commands

The user can also enable the scan test mode with the button "Scan Test" on the toolbar and with the command "Scan Test" under the "Options" menu.

The user can enable the greyscale image window by right-clicking the IRISPen control on the Windows taskbar and selecting the command "Greyscale Image" under the "Options" menu, and with the command "Greyscale Image" under the "Options" menu.

Help Topics

Allows to navigate through the on-line help.

Refer to the command "Help Topics" under the "Help" menu for further information.

Related Command

The user can also access the on-line help system with the command "Help Topics" under the "Help" menu.

Exit

Closes the IRISPen software.

Refer to the command "Exit" under the "File" menu for further information.

Related Command

The user can also quit the IRISPen application with the command "Exit" under the "File" menu.

GREYSCALE IMAGE WINDOW

The greyscale image window displays the "raw" greyscale images as they are generated by the pen scanner.

Scanning in greyscale is in some instances necessary to obtain good OCR results. When text is printed on a color background, scanning in color may create the tone differences that are lacking in black-and-white images. When there is only limited contrast between the text and the background, the background can create "noise" that renders the recognition difficult or impossible! Think for instance of black text printed on a dark background: when scanning such a document in black-and-white, the user would not be able to "drop" the background color without losing the text information as well.

Related Commands

The user can enable the greyscale image window by right-clicking the IRISPen control and selecting the command "Greyscale Image", and by selecting the command "Greyscale Image" under the "Options" menu.

The scan test mode displays a black-and-white version of the greyscale images generated by the pen scanner alongside the reading results, and allows advanced image adjustment to improve the reading results. The user can enable the scan test mode by right-clicking the IRISPen control on the Windows taskbar and selecting the command "Scan Test", by clicking the "Scan Test" button on the toolbar and by selecting the command "Scan Test" under the "Options" menu.

STATUS WINDOW

This window displays multiple system information useful to the user and provides direct access to the IRISPen application window.

The user can drag the status window to any appropriate position on the screen.

Alphanumeric-Bar Codes, English, Dot Matrix, Application , Left-Handed, Negative

Besides the "basic" settings such as reading mode, language and output target, the status window will also tell you if bar code reading is enabled as secondary reading mode, if the dot matrix mode is enabled, if left-handed scanning is used and if light text on a dark background is scanned.

For the IRISPen Translator, the status window lists the translation direction. When English text is translated into French, "English-French" is mentioned in the status window.

Alphanumeric, English>French, Application

When the IRISPen is used to scan small graphics (by selecting the reading mode "Black-and-White Bitmap to Clipboard" or "Greyscale Bitmap to Clipboard" in the IRISPen application window), a blinking status window displays the message "Image in Clipboard" to indicate that the graphic is in the Windows clipboard, ready to be pasted in the target applications. (The user can paste the graphic immediately, he does not have to wait until this message disappears.)

Image in clipboard

When the multiline scanning mode is used (by enabling the option "Multilines" in the IRISPen application window, a blinking status window displays the message "Press pen button to flush" for a few seconds after each scan. (The user can rescan immediately, he does not have to wait until this message disappears.)

Click pen button to flush

When the user scans too fast, a blinking status window asks him to scan slower for a few seconds. (He can scan some 3.15" (8 cm.) per second.)

Maximal scanning speed exceeded: please scan slower

The user can right-click or double-click the status window to restore the IRISPen application window.

Related Commands

The user can enable the status window by right-clicking the IRISPen control on the Windows taskbar and selecting the command "Status Window", by selecting the command "Status Window" from the "Options" menu and by loading IRISPen settings with the command "Open Settings" under the "File" menu.

The user can right-click the IRISPen control on the Windows taskbar and select the command "Open" to restore the IRISPen application window. He can also define a "restore" shortcut with the command "Hotkeys" under the "Settings" menu.

TRANSLATION WINDOW (IRISPEN TRANSLATOR)

A "pop-up" speech bubble called translation window displays the sentence translation generated by the IRISPen Translator.
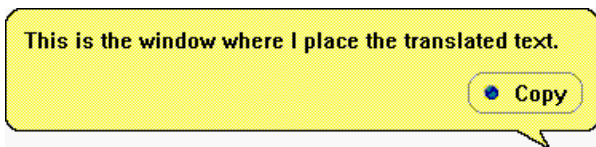
The size of the speech bubble always fits the amounts of text to be displayed on-screen. When large amounts of texts are submitted to the translation module, the translation window may display some 2,000 characters. The entire screen will be occupied, "..." is added at the end to indicate that the text was truncated.

Unrecognized characters are by default represented by a tilde (the "~" symbol), words unknown to the system are repeated literally.

The translation window appears when the first sentence is translated and stays there until a new sentence is translated. The user has to quit the IRISPen Translator software, disable the translation function by selecting the option "<None>" with the option "Translate to" in the IRISPen application window or

click the "Copy" button of the translation window to make it disappear. The "Copy" button copies the translation to the clipboard: the translation window disappears as there is no further need to display it on-screen.

The "Copy" button allows the user to transfer the translated text to other windows via the clipboard. The user can paste the translation in the same window as the one where the scanned text is placed to study the foreign sentences and their meaning comfortably later on. He can also insert the translation in a different window to save the original text and its translation in separate files.

When the clipboard translation is enabled with the option "Translate Clipboard" under the "Options" menu, the translation window displays the "Translate Clipboard?" button as soon as text is copied to the clipboard. (Graphics that are copied to the clipboard are simply ignored.) This button is clicked to generate a translation for the contents of the clipboard.

The IRISPen Translator has sufficient linguistic intelligence to translate abbreviations, normalized character strings such as "1st" and phrases consisting of several words such as "state of mind" (état d'esprit), "take away" (wegnehmen) etc.

However, the user should be aware that computer-aided translation gives an imperfect first draft, an "information-only" translation. Should a translation be published, it will need to be refined by a human translator. The IRISPen Translator should be used to translate business-like, non-emotive language as computers do not cope well with connotation, metaphor and irony. Literature, poetry and some advertizing literature do not translate well, and often lead to humorous effects.

TOOLBAR

The toolbar provides immediate access to all frequently used commands.

Two buttons - "Bar Codes" and "Handprinting" - are limited to the IRISPen Executive; the "Speech" button is common to the IRISPen Executive and IRISPen Translator.

Scan Test

Enables the scan test mode where the user displays a black-and-white version of the greyscale images generated by the pen scanner and the reading results.
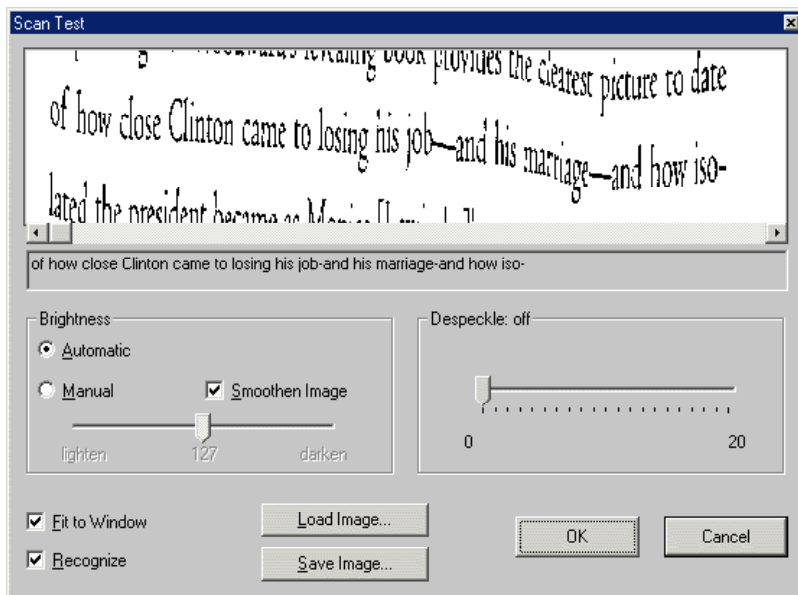
Scan tests are used to get acquainted with the operation of the IRISPen. The I.R.I.S. test sheets can be used to this end.

Furthermore, the scan test window is equipped with advanced imaging features to adjust the scanning to the color tones of the document.

Scanning in greyscale is in some instances necessary to obtain good OCR results. When text is printed on a color background, scanning in color may create the tone differences that are lacking in black-and-white images. When there is only limited contrast between the text and the background, the background can create "noise" that renders the recognition difficult or impossible! Think for instance of black text printed on a dark background: when scanning such a document in black-and-white, the user would not be able to "drop" the background color without losing the text information as well.

Powerful intelligent routines automatically create a black-and-white version for every greyscale image. Should this be necessary, the user can optimize the image further for the consecutive OCR process with this command. In doing so, a much larger range of documents becomes legible.

The recognition is re-executed promptly each time these settings are modified. The user is recommended to keep the greyscale image visible when he's adapting these parameters.

The option "Smoothen Image" renders the greyscale images more homogeneous by "flattening", smoothing out relative differences in intensity. As a result, a sharper contrast is created between the foreground - the text - and the background - a color, artwork etc. It is recommended that this option be enabled at all times.

(A red or yellow background will never be picked up by the pen scanner: red and yellow are the IRISPen's "drop out" or invisible colors. This implies that all text printed in red and yellow is illegible.)

The option "Brightness" influences the black-and-white threshold. The setting "Automatic" determines the bilevel threshold automatically but with a great deal of intelligence. The brightness is set automatically for all "normal" documents. A different threshold is applied when necessary by darkening or lightening the black-and-white image: when the image is darkened, more pixels become black in the black-and-white version, when the image is lightened, less pixels become black.

The option "Despeckle" concerns the black-and-white image. "Despeckling" means that the "parasite pixels" (also called "salt and pepper noise") will be removed from the black-and-white images. The user shouldn't erase spots that are too big, otherwise he runs the risk of erasing the dots on "i" etc., portions of dot matrix letters etc.!

Clicking the button "Load Image" allows to open and recognize prescanned images. This is purely a diagnostic feature.

Clicking the button "Save Image" saves the scanned image in a graphic file. Three file formats are available: greyscale and black-and-white Windows bitmaps (BMP) and a proprietary black-and-white graphic format IRISPen Graphic. IRISPen graphics should be used for diagnostic purposes only, bitmap files can also be used to save small images - logos, signatures, mathematical formulae etc. Thanks to the scan test mode, the user can decide whether an image is suitable to be saved. He can scan several times and save the scanned image which he likes most because it isn't skewed, distorted etc.

The option "Fit to Window" ensures that the entire image is displayed in the scan test window, whatever the length of the scanned line may be. (The line length is limited to 12" (30.5 cm.) per scan.) With this option disabled, the image isn't compressed horizontally and the scroll bar is used to display various parts of the image successively.

When the scanned line is very long, the user may have to scroll to display the recognized text entirely. He places the cursor in the text zone and presses the End key to go directly to the end of the recognized text; he presses Home to return to the beginning.

The option "Recognize" sees to it that all scanned or opened images are recognized. When the scan test mode is used to save graphics instead of scanning data, this option can be disabled: scanning a graphic, the IRISPen software won't find any text anyway. (When the reading mode "Black-and-White Bitmap to Clipboard" or "Greyscale Bitmap to Clipboard" is selected in the IRISPen application window, the option "Recognize" is replaced by the option "Bitmap to Clipboard".)

Related Commands

The user can also enable the scan test mode by right-clicking the IRISPen control on the Windows toolbar and selecting the command "Scan Test", and by selecting the command "Scan Test" under the "Options" menu.

The user can enable the greyscale image window by right-clicking the IRISPen control on the Windows taskbar and selecting the command "Greyscale Image", and by selecting the command "Greyscale Image" under the "Options" menu.

Pen Button

Allows the programming of the single and double click of the pen scanner button.

The single and double click of the pen scanner's button can be programmed to execute any combination of characters and keyboard commands and to switch from one reading mode to another.

As is the case with a computer mouse, the single and double click of the pen scanner button are programmed separately.
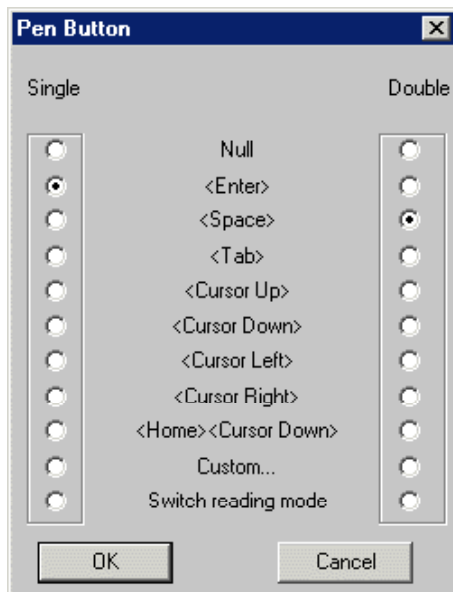
Some frequent options such as <Tab>, <Enter>, <Space> etc. are presented by default, but with the option "Custom", the user can program any keyboard command to be executed or character string to be inserted when the pen button is pressed. The option "Null" is selected to disable the applicable pen button command.

When reading lines of text, the pen button can be programmed to execute an <Enter> command. Pressing the button ensures that a new paragraph is started.

The pen button can also be used to insert text which is not on the paper. If a price list is read where the currency is missing once in a while, the pen button can be used to insert "$" where necessary.

Furthermore, the pen button makes it easier to insert the scanned text at the right location. When data is entered in a database, the <Tab> key can be assigned to the pen button to move the cursor to the next field.

One application is encoding addresses from business cards, brochures, catalogs etc. collected during a trade show in a database. Depending on the addresses, some fields will be skipped. For some addresses, the country will be added, and for others, it won't, some contacts have an e-mail address while others don't etc. To encode these data efficiently, the single pen button click is programmed as <Tab>, even though <Tab> already functions as ending code. The suffix <Tab> moves the cursor to the next field automatically, and the pen button is clicked once when a database field should be skipped. To start a new record when an address was completed, the user normally clicks a button with his mouse, but given the pen button, he can program the double click to correspond to that button's shortcut.

To encode forms with an address (alphanumeric data) and coded numbers (numeric data) where the data must be distributed among different database fields, the single button can be used to move the cursor to a new database field and the double button click to toggle between the alphanumeric and numeric reading mode.

To encode data in a spreadsheet where no vertical bars separate the columns, the figures are read one by one and the single click is programmed as <Tab> to move the cursor to the next cell, the double click is programmed as <Home><Cursor Down> to move the cursor to the next row.

When data of varying nature - printed or typed text, printed and handprinted numerals - is mixed on the page, the pen button can be programmed to switch reading modes. When the user has selected "Alphanumeric" as his reading mode and he clicks the pen button, the IRISPen switches to the mode "Bar Codes". When he clicks a few times more, he switches to "Black-and-White Bitmap to Clipboard", "Greyscale Bitmap to Clipboard", "Handprinting", "Numeric" and back to "Alphanumeric".

(Switching from one reading mode to another is superfluous to read bar codes alongside data of another type: the option "Bar Codes" in the IRISPen application

window enables a mixed reading mode where bar codes are read alongside "normal" data. The parameter "Reading Mode" in the IRISPen application window determines what kind of data is read, this option determines whether read bar codes can be read at the same time. The mixed reading mode is limited to the IRISPen Executive.)

When multiline scanning is enabled with the option "Multiline" in the IRISPen application window, the pen button takes on a special meaning. Multiple lines are "buffered" before they are inserted, spoken and translated. (This option is limited to the IRISPen Executive and the IRISPen Translator.) It takes a click of the pen button to "flush" the scanned text, to insert it in the active application and start the speech and translation process.

It should be noted that the pen button remains a programmable command button, even when the "multiline" mode of operation is enabled. When the multiline scanning is enabled, the first click of the pen button triggers the "flush", all consecutive pen button clicks are interpreted according to their programming (if any). An immediate *double*-click preceding the flush is also interpreted as the flush command.

Imagine that the single button click is programmed to execute <Tab> and the double button is programmed to click to insert "$". After each scanned line, the status window blinks with the message "Press pen button to flush". As soon as the text is "flushed" with the first pen button click, an additional pen button click executes the <Tab> key to move the cursor to the next field. The double button click inserts "$" in the application.

Related Commands

Pen button programming can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

To execute a keyboard command or insert text *fully* automatically, the optioons "Starting Code" and "Ending Code" of the "Alignment" button on the toolbar is used.

The user can program a hotkey to go through the reading modes from within the active application with the command "Hotkeys" under the "Settings" menu.

Alignment

Indicates how the scanned line of text is interpreted. Is cropping applied to the recognized text? Is data or is a keyboard command to be added in front and behind?

The cropping options can be used in several situations. First of all, to isolate keywords from their context. If only specific words should be entered in a database, the user can start scanning somewhere in the word in front of and end scanning in the word after the word to be captured. The word cropping sees to it that only the intended, required words are passed on.

Another example: when reading text set in columns with a small gutter, it is not always easy to isolate a column. The word cropping allows to scan part of the neighboring columns without generating unwanted output.

A minor option, "Ignore Rejected Characters in First and Last Position", concerns unrecognized characters: unknown symbols at the beginning and end can be ignored. These are usually the result of partial scanning: the user has placed his IRISPen in the middle of these characters, and the system is therefore unable to recognize them correctly.

The options "Starting Code" and "Ending Code" will add text or a cursor command to the recognized text. A starting code is any combination of characters and keyboard commands added in front of the recognized text, an ending code is the same added at the end of the recognized text.

When scanning a few lines of text, the user normally selects the <Space> key as suffix to add a space after each line. When reading two lines, "The second line" and "follows the first.", the resulting output becomes "The second line follows the first.". The space after the word "line" is added by the IRISPen software.

When translating text with the IRISPen Translator, <Space> should be selected as ending code to ensure that a space is added after each scanned line automatically, otherwise the words "line" and "follows" would be glued with the intranslatable sentence "The second **linefollows** the first." as a result.

When scanning text in a database, the user chooses the cursor command <Tab> as suffix to reposition the cursor automatically in the next field, ready for the next scan.

Other typical applications where starting and ending codes are used to add text are encoding price lists where the currency is added, for example "$" or "USD", to the recognized figures, and encoding telephone numbers from a list and adding "Tel: " in front.

To add text and execute keyboard commands at the user's command, the pen button is programmed with the "Pen Button" button on the toolbar.

The option "Restore Paragraph" restores paragraphs by "recomposing" hyphenated words and adding suffixes only where they are needed. When this

option is enabled, end hyphens are deleted and the ending code - if one is defined - is ignored. When two lines are read, "The new presi-" and "dent waved from the balcony.", the resulting output is "The new president waved from the balcony.". The end hyphen of "presi-" is ignored and so is the "normal" suffix <Space>. If the two following lines are "His lovely wife" and "had joined him.", the output will be "His lovely wife had joined him.", with the space after the word "wife" added as suffix by the IRISPen software.

This option is mandatory for the IRISPen Translator to avoid ungrammatical and therefore intranslatable text: it restores the sentences by "recomposing" hyphenated words and adding the suffix <Space> only where it is needed. (The IRISPen Translator does more here than recognize text: the recognized text is contextually adjusted, or "pre-edited" as other translation solutions would call it, to allow correct translation.)

Related Command

The alignment options can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

Output Subset

Allows to define customized character sets.

This command complements the alphanumeric and numeric character modes as selected in the IRISPen application window. When the alphanumeric reading mode is active, the set of recognized symbols is obviously influenced by the selected language.
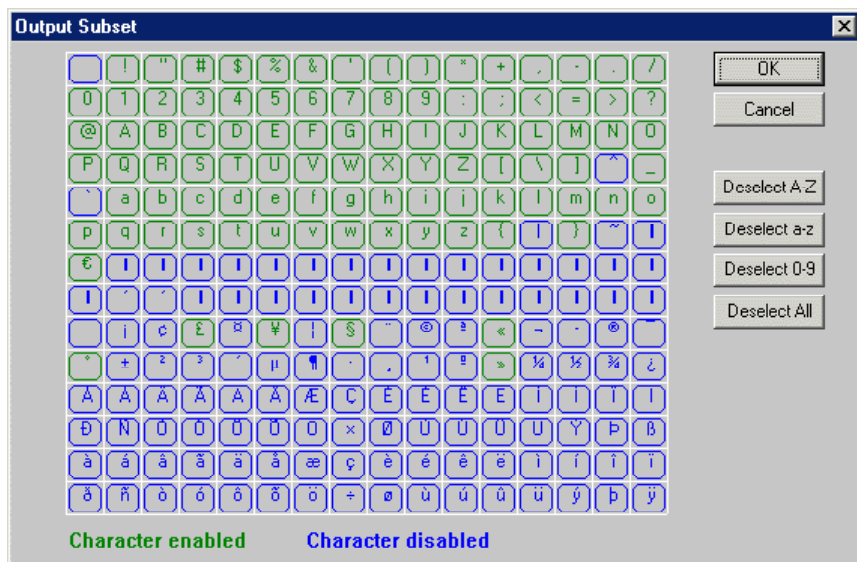
When a limited number of symbols should be recognized, the IRISPen's accuracy and speed can be improved by narrowing down the output subset. The user could for instance enter serial numbers and product codes where the numerals are completed by a few alphabetic symbols in a database.

This is all the more important as OCR may generate recognition errors for those symbols with very similar characteristics. When errors do occur, they usually concern "B" and '8', "O" and '0', "l" and '1' etc.

The user adds and removes symbols by clicking them. Some buttons are available to add and remove entire ranges of characters simultaneously: "Select All", "Select 0-9" etc.

A concrete example is creating an OCR-B character set which limits recognition to the numerals and the symbols "<", ">" and "+". (OCR-B is at the same time a highly specialized typeface used only in financial documents such as printed

checks, bank statements etc. I.R.I.S. has a specialized version of the IRISPen for the financial market.)

It should be stressed that the IRISPen does not recognize all characters as are displayed in this window. Some special characters such as ¥ and © are *not* recognized by the IRISPen but can be used in character and table frame conversions, in starting and ending codes and be inserted by the pen button clicks.

The window below indicates which symbols can and can not be recognized - the highlighted symbols are *not* recognized by the IRISPen.
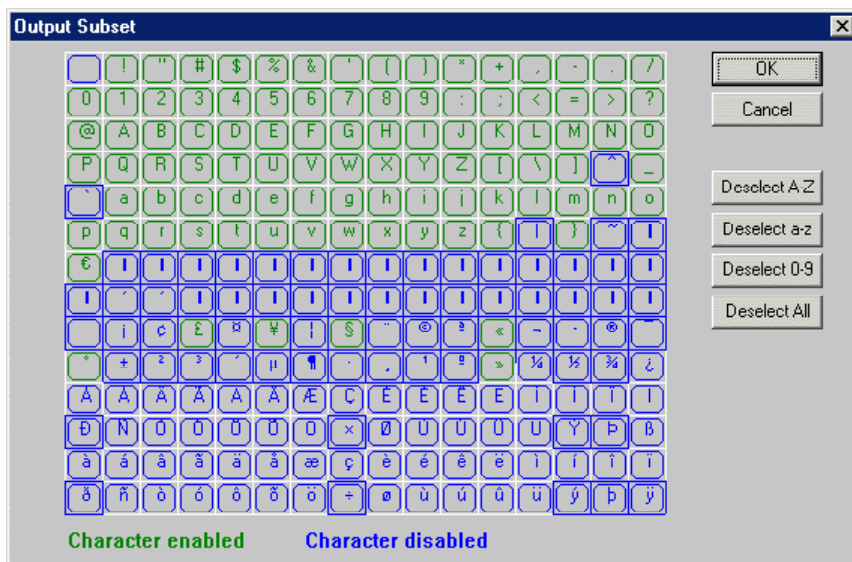
Related Command

Output subsets can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

Bar Codes (IRISPen Executive)

Enables the bar code reading mode and defines the bar code settings.

The bar code reading mode can also be selected in the IRISPen application window. When *this* button is clicked, the user is simultaneously given access to the bar code settings.
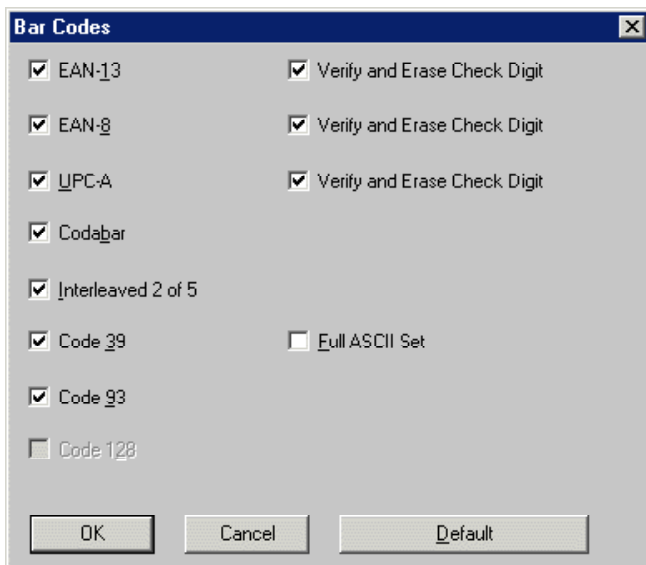
The user is given full control over the used symbologies. By default, all bar code symbologies are enabled and the IRISPen will autodetect the used symbology. The user does not have to adjust the bar code settings to scan bar codes. Contrary to most bar code readers, the IRISPen software detects fully automatically which bar code type is used. Various bar codes can be mixed without any consequence.

Refer to chapter 4 for scanning tips and design rules for forms and labels; refer to appendix B for a detailed discussion of the supported bar code symbologies.

The user can limit the recognition to specific bar code types by clicking the corresponding options. The "Default" button restores the default bar code settings.

Additional options are available for the symbologies EAN-13, EAN-8, UPC-A and Code 39.

For the symbologies EAN-13, EAN-8 and UPC-A, the user can verify the check character with the option "Verify and Erase Check Digit". When this option is enabled, the check digit is verified but filtered from the output. When this option is disabled, the check character is not verified but read and included in the output. This option should be disabled to read Code 39 bar codes which do not contain a check digit.

The option "Full ASCII Set" concerns the Code 39 symbology. An extended version of Code 39, a symbology which in its standard version has a limited character set, allows the encoding of all ASCII characters. The "full" ASCII set refers to all characters up to character 128 in the ASCII table set, the "special" and accentuated symbols from 128 onwards such as ç and ó are *not* included. Enabling this option allows the interpretation of that symbology. The full ASCII symbol set is not supported by default, as it is a special version of the Code 39 symbology.

Related Commands

The reading mode "Bar Codes" in the IRISPen application window enables bar code reading as *only* reading mode.

The option "Bar Codes" in the IRISPen application window enables a mixed reading mode where bar codes are read alongside data of another type. The parameter "Reading Mode" in the IRISPen application window then determines the primary reading mode and the option "Bar Codes" in the IRISPen application window determines whether read bar codes can be read simultaneously.

The "hands- and eyes-busy" activity of reading bar codes can be combined with auditive feedback by enabling the "Speech" button on the toolbar.

Bar code settings are included in the IRISPen settings. The command "Save Settings" under the "File" menu and terminating the software with the option "Save Settings on Exit" enabled, include the selected values in the IRISPen settings.

Handprinting (IRISPen Executive)

Enables the handprinting reading mode and defines the handprinting settings.

The handprinting reading mode can also be selected in the IRISPen application window. When *this* button is clicked, the user is simultaneously given access to the handprinting settings.

The user is given full control over spacing and rejection levels.
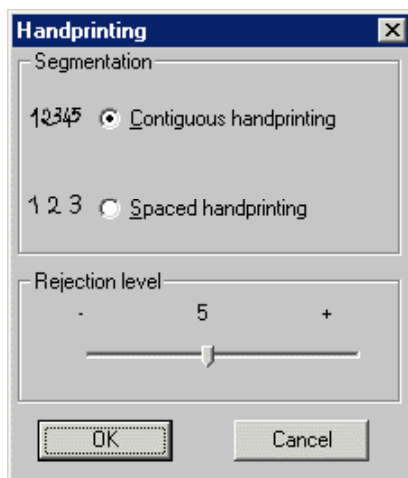
Segmentation determines whether the characters are well spaced or contiguous. With spaced handprinting, each character is distinctly separated from the

others. With contiguous spacing, the space between two characters is considerably reduced. Some characters may occassionally touch each other and the zones covered by a single character may overlap.

It is mandatory to set this parameter correctly. When contiguous handprinting is read with the segmentation set to spaced, some characters may be skipped by the system. With contiguous printing, broken characters may be misread as double characters.

An easy way of guaranteeing spaced printing is the use of forms where the individual numerals must be placed in predefined blocks. To ensure that these predefines zones are not picked up by the IRISPen, the invisible or "drop out" colors red and yellow should be used.

A slider bar allows to adjust the rejection level. When the rejection threshold is increased, characters will result in unknown symbols quicker, there will be less substitutions; when the rejection rate is decreased, the system will substitute more symbols as it tries harder to find a match.

For applications where extreme accuracy is essential, a high rejection level should be maintained. When structured character strings such as client numbers, product codes etc. are encoded, it is useful to incorporate a check digit which is verified by the encoding utility.

Handprinting recognition is limited to the numerals and the symbols , (comma), . (dot) and - (hyphen).

When handprinted and other data are mixed on the page, the user can switch reading modes by defining a hotkey with the command "Hotkeys" under the "Settings" menu or program the pen button to this end with the "Pen Button" button on the toolbar.

The "hands- and eyes-busy" activity of reading handprinting can be combined with auditive feedback by enabling the "Speech" button on the toolbar. The speech feature incorporates intelligent routines that allow it to recognize telephone numbers, bank account and social security numbers, dates etc. When bank account numbers are read, short pauses are inserted where a dash occurs, when tele-

phone numbers are read, the speech synthesis will add the word "area code" where it is useful etc.

Refer to appendix C for scanning tips and writing and form design rules.

Speech (IRISPen Executive - IRISPen Translator)

Activates the "Read & Speak" speech synthesis.

With the IRISPen Executive, speech synthesis provides clear, natural sounding auditive feedback of the recognized data. "Hands- and eyes-busy" data entry becomes possible as the user no longer has to look up from his documents to verify the results on-screen. Text-to-speech also opens a world of multimedia applications to the IRISPen users.

With the IRISPen Translator, speech synthesis provides clear, natural sounding pronunciation of the scanned source text. To learn how to speak a language, the user can just scan a text and hear the PC pronounce the words. (To pronounce utterances from the other language, the user inverts the source and target languages in the IRISPen application window.)

The user's PC must be equipped with a Windows compatible sound card if speech synthesis is to work.

Speech synthesis should not be enabled unless the user intends to use it: the recognition slows down as the user waits for the recognized data to be read aloud.

The user is warned when he tries to read a language for which no speech module is available.

Enabling the speech feature disables the option "Sound" on the toolbar automatically: mixing spoken text and system sounds makes no sense.

For optimal speech results, the user is recommended to enable the multiline scanning mode with the option "Multilines" in the IRISPen application window. Multiline scanning ensures a natural prosody: when the text-to-speech module has to pronounce individual lines - "chopped up" text bits - rather than full sentences, the intonation may be somewhat less than "human".

The user can repeat the pronunciation of the last scan (or "flush" when the multiline mode is enabled) by *right*-clicking the "Speech" button and by selecting the command "Repeat Speech" under the "Options" menu. The user can also create a hotkey for the speech repetition with the command "Hotkeys" under the "Settings" menu.

The speech synthesis capability is limited to the IRISPen Executive - where the user is entitled to a single language of his choice - and to the IRISPen Translator - where the speech feature is offered for the language pair. The English-French version of the IRISPen Translator for instance has the English and French speech libraries.

Speech modules for additional languages can further extend his system. Ten languages are currently available: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.

Related Command

The speech feature can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

Sound

This option enables the system sound.

When it is enabled, the system generates sound when the IRISPen starts scanning by pressing the pen scanner onto the paper, stops scanning by lifting the pen scanner away from the paper and when the recognized text or scanned graphic is sent to the target application or clipboard.

Enabling the speech feature disables this option automatically: mixing spoken text and system sounds makes no sense.

Related Command

This option can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

*U*SER'S *G*UIDE

FILE MENU

The "File" menu has a limited function: it lets the user open and save IRISPen settings and quit the software. The last four settings are also available in the "File" menu.

Open Settings

Loads IRISPen configurations.

Settings files contain all operational parameters: the reading mode, the document language, the target language for the translation (IRISPen Translator), whether dot matrix reading mode (IRISPen Executive) and the mixed reading mode with secondary bar code reading is enabled (IRISPen Executive), the output target, whether the multiline reading mode is enabled (IRISPen Executive and IRISPen Translator), the programming of pen button clicks and hotkeys,

table frame and character conversions, the output subset, alignment options (including starting and ending codes), the bar code and handprinting settings (IRISPen Executve), whether the speech synthesis (IRISPen Executive and IRISPen Translator) or system sound is enabled, whether the clipboard translation (IRISPen Translator), the left-handed scanning and the image inversion are enabled.

Related Commands

The last four settings are available in the "File" menu; clicking the files suffices to load them.

To load the default settings, the user selects the file DEFAULT.IPC. (When this file is missing, the factory settings apply.)

Settings are saved with the command "Save Settings" under the "File" menu.

Save Settings

Saves the current IRISPen settings in a configuration file for future use.

Settings files contain all operational parameters: the reading mode, the document language, the target language for the translation (IRISPen Translator), whether dot matrix reading mode (IRISPen Executive) and the mixed reading mode with secondary bar code reading is enabled (IRISPen Executive), the output target, whether the multiline reading mode is enabled (IRISPen Executive and IRISPen Translator), the programming of pen button clicks and hotkeys, table frame and character conversions, the output subset, alignment options (including starting and ending codes), the bar code and handprinting settings (IRISPen Executve), whether the speech synthesis (IRISPen Executive and IRISPen Translator) or system sound is enabled, whether the clipboard translation (IRISPen Translator), the left-handed scanning and the image inversion are enabled.

This command is primarily used to save specific IRISPen settings. However, it also allows the user to modify the default configuration. To do this, the user saves the current settings in the file DEFAULT.IPC.

Related Commands

An easier way of saving the current settings as default is enabling the option "Save Settings on Exit" under the "Options" menu before the software is terminated. These settings will be used at the next startup.

Settings are loaded into memory with the command "Open Settings" under the "File" menu.

Exit

Closes the IRISPen software.

As the user quits, he also saves the current settings as default settings, provided that the option "Save Settings on Exit" under the "Options" menu is enabled. If this is the case, these settings will be used at the next startup. The default settings are saved in the file DEFAULT.IPC. When this file is erased, the factory settings apply.

Settings files contain all operational parameters: the reading mode, the document language, the target language for the translation (IRISPen Translator), whether dot matrix reading mode (IRISPen Executive) and the mixed reading mode with secondary bar code reading is enabled (IRISPen Executive), the output target, whether the multiline reading mode is enabled (IRISPen Executive and IRISPen Translator), the programming of pen button clicks and hotkeys, table frame and character conversions, the output subset, alignment options (including starting and ending codes), the bar code and handprinting settings (IRISPen Executve), whether the speech synthesis (IRISPen Executive and IRISPen Translator) or system sound is enabled, whether the clipboard translation (IRISPen Translator), the left-handed scanning and the image inversion are enabled.

Related Commands

The user can also quit the IRISPen application by right-clicking the IRISPen control on the Windows taskbar and selecting the command "Exit", and by clicking the close button of the IRISPen application window.

Last Settings

The four settings last used are displayed under the "Exit" command. Clicking on the files suffices to load the settings.

Settings files contain all operational parameters: the reading mode, the document language, the target language for the translation (IRISPen Translator), whether dot matrix reading mode (IRISPen Executive) and the mixed reading mode with secondary bar code reading is enabled (IRISPen Executive), the output target, whether the multiline reading mode is enabled (IRISPen Executive and IRISPen Translator), the programming of pen button clicks and hotkeys, table frame and character conversions, the output subset, alignment options (including starting and ending codes), the bar code and handprinting settings (IRISPen Executve), whether the speech synthesis (IRISPen Executive and IRISPen Translator) or system sound is enabled, whether the clipboard translation (IRISPen Translator), the left-handed scanning and the image inversion are enabled.

Related Command

The command "Open Settings" under the "File" menu allows to load settings files into memory.

SETTINGS MENU

The "Settings" menu contains many advanced features - alignment, programming of pen button clicks, character and table frame conversion and customized character sets - that fine-tune the operation of the IRISPen to the user's application.

All text-related commands - alignment, character and table frame conversions, customized character sets - are disabled when the reading modes "Bar Codes" and "Handprinting" are active.

Hotkeys

Allows the user to program shortcut keys for two common operations: restoring the IRISPen application quickly to the foreground and switching between the reading modes from within the target applications. For the IRISPen Executive and IRISPen Translator, a third function can be assigned to a hotkey - the repetition of the utterance last spoken.

Key combinations such as Ctrl+V used by the target application(s) should be avoided for obvious reasons.
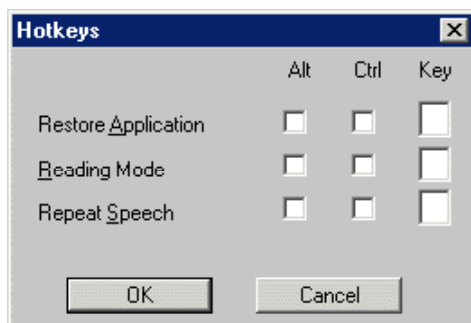
As the IRISPen application normally runs in the background, the option "Restore Application" defines a shortcut key for restoring the IRISPen application window from within the active applications.

Alternatively, the user can right-click or double-click the status window to "restore" the IRISPen application; he can also right-click the IRISPen control on the Windows taskbar and select the command "Open".

The second hotkey is used to switch between the available reading modes. When data of varying nature - printed or typed text, printed and handprinted numerals - are mixed on the page, a hotkey can be programmed to switch reading

modes from within the active application. (To combine a reading mode with bar code reading, a mixed reading mode can be enabled with the option "Bar Codes" in the IRISPen application window.)

When the user has selected "Alphanumeric" as his reading mode and he clicks the pen button, the IRISPen switches to the mode "Bar Codes". When he clicks a few times more, he switches to "Black-and-White Bitmap to Clipboard", "Greyscale Bitmap to Clipboard", "Handprinting", "Numeric" and back to "Alphanumeric". Possible applications are documents where different data types are mixed on the page.

For the IRISPen Executive and the IRISPen Translator, a hotkey is available for the speech repetition. Pressing the programmed hotkey from within the active application "replays" the pronunciation of the last sentence or paragraph.

Related Commands

Hotkey programming can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

The user can program the pen button to switch from one reading mode to another with the "Pen Button" button on the toolbar.

The parameter "Reading Mode" in the IRISPen application window determines what kind of data is read, the option "Bar Codes" determine whether read bar codes are read at the same time. The user then enables a mixed reading mode where bar codes are read alongside data of another type. (This option is limited to the IRISPen Executive.)

Speech repetition is also available by either *right*-clicking the "Speech" button or by clicking the command "Repeat Speech" under the "Options" menu. (This command is limited to the IRISPen Executive and the IRISPen Translator.)

Pen Button

Refer to the "Pen Button" button on the toolbar.

Alignment

Refer to the "Alignment" button on the toolbar.

Table Frame Conversion

Allows the interpretation of the table frame that separates table cells.

Some common options are presented by default, but the user can define any combination of character and keyboard commands.
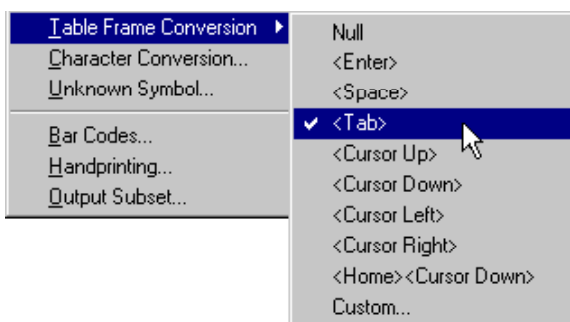
When the user selects <Tab>, the IRISPen software will execute a <Tab> whenever a vertical bar is encountered. As a <Tab> moves the cursor to the next cell, scanning a table into a spreadsheet will put the individual figures in separate cells. In this way, the typical table structure with rows and columns is recreated by simply scanning each line in turn.

This feature only works when continuous vertical bars separate the table columns: the IRISPen software has to encounter a physical mark, detect a border between two columns.

Where the original document does *not* contain vertical bars, the easiest procedure is to quickly add them with a pen and a ruler. In some cases, the "Character Conversion" command offers the answer as it allows replacement of scanned characters by a cursor command. For example, to insert the line "2,55$ 5.99$ 65$" in a worksheet, it suffices to replace the symbol "$" by <Tab> to place the individual figures in separate cells.

Related Command

Table frame conversion can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

Character Conversion

Allows the replacement and erasure of recognized characters.

The IRISPen not only recognizes data but also replaces them: the user can recognize a character and replace it by any combination of characters and keyboard commands. He can even delete a character (replace it by nothing).
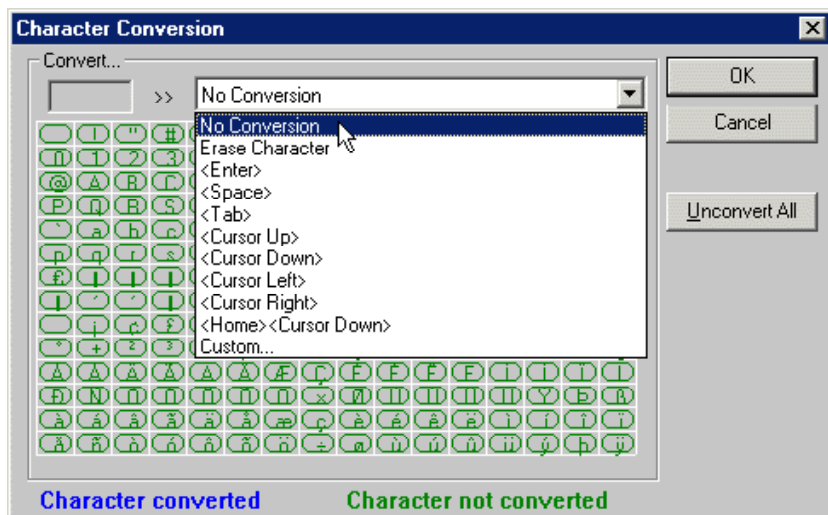
Characters to which conversion apply are highlighted. When the user clicks on them, the applicable conversion is displayed and he can define, modify or cancel the character conversion. The option "No Conversion" is selected to disable the applicable conversion, the button "Unconvert All" disables all active conversions.

A realistic example is scanning a price list mentioning the currency into a spreadsheet. As the currency is of no interest, the "$" symbol is erased to filter the currency from the output.

Related Commands

End hyphens in body text can be erased automatically when the option "Restore Paragraph" is enabled. Refer to the "Alignment" button on the toolbar.

Character conversions can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

Unknown Symbol

Determines by which symbol a rejected, unrecognized character is represented. The default symbol is a tilde (the "~" symbol).
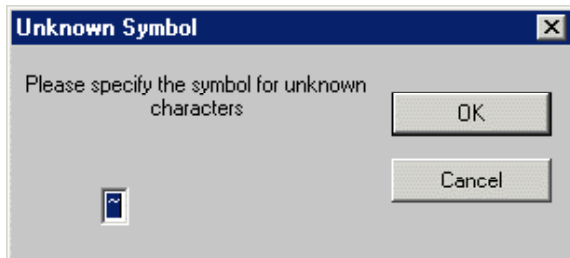
Words unknown to the translation system are repeated literally in the translation window of the IRISPen Translator.

Related Command

Rejected symbols at the beginning and end of a scan are ignored automatically when the option "Ignore Rejected Characters in First and Last Position" is enabled. Refer to the "Alignment" button on the toolbar.

Bar Codes (IRISPen Executive)

Refer to the "Bar Codes" button on the toolbar.

<u>Handprinting (IRISPen Executive)</u>

Refer to the "Handprinting" button on the toolbar.

<u>Output Subset</u>

Refer to the "Output Subset" button on the toolbar. Refer to the "Output Subset" button on the toolbar.

OPTIONS MENU

The "Options" menu enables the scan test mode and contains some optional system settings. With the IRISPen Translator, this menu also contains the speech repetition command and enables the clipboard translation mode.

<u>Scan Test</u>

Refer to the "Scan Test" button on the toolbar.

<u>Speech (IRISPen Executive and IRISPen Translator)</u>

Refer to the "Speech" button on the toolbar.

<u>Repeat Speech (IRISPen Executive and IRISPen Translator)</u>

Repeats the pronunciation of the last line (or "flush" when the multiline scanning mode is enabled with the option "Multilines" of the IRISPen application window).

The pronunciation of the last uttereance can also be repeated by either *right*-clicking the "Speech" button or by creating a speech repetition keyboard shortcut with the command "Hotkeys" under the "Settings" menu.

<u>Sound</u>

Refer to the "Sound" button on the toolbar.

Left-Handed Scanning

This option adjusts the scanning direction for left-handed users.

By default, the pen scanner is operated from left to right, but with this option enabled, the pen scanner is ruled over the text from right to left. This only works when the pen scanner is held rotated by 180° so that the pen button faces towards the user, while the document itself retains its normal orientation. The software "recomposes" a correct image by rotating it by 180°.

Alternatively, when the page *and* the pen scanner are rotated by 180°, the option "Left-Handed Scanning" can be disabled but the user can still scan right-to-left.

The enabling of left-handed scanning is indicated in the status window. It is disabled by default.

Related Command

The right-to-left scanning direction can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

Negative Image

Inverses the scanned image.

The image in the greyscale image window is left untouched. The inversion is only visible in the black-and-white image: black "pixels" become white, white pixels become black. As a result, light data on a dark background becomes legible.

This scanning mode can also be used to generate "negative", inverted graphics (sent to the clipboard or saved in a graphic file).

The enabling of image inversion is indicated in the status window.

Related Command

The image inversion can also be enabled by loading IRISPen settings with the command "Open Settings" under the "File" menu.

Translate Clipboard (IRISPen Translator)

Enables the clipboard translation mode.

When clipboard translation is enabled, the user is prompted to translate all text copied to the Windows clipboard. The translation window displays the button "Translate Clipboard?" as soon as text is copied to the clipboard. (Graphics copied to the clipboard are simply ignored.) The user clicks this button to generate a translation for the contents of the clipboard.

The capability to translate any text sent to the clipboard has far-reaching implications. It allows the user to translate *any* text he highlights on his screen and copies to the clipboard: text entered on the keyboard, wordprocessor files, mail messages received in a foreign language, Internet pages displayed with a web browser etc.

The user can also send the scanned text directly to the clipboard - the recognized text will *not* be inserted in the active application - by selecting the clipboard as output target with the setting "Output Target" in the IRISPen Translator application window. The text to be translated is scanned, but the source text is not retrieved as only its translation is considered relevant.

Related Command

The setting "Language" in the application window of the IRISPen Translator indicates the source language, the setting "Translate to" in the application window of the IRISPen Translator indicates the target language.

Greyscale Image

Refer to the command "Greyscale Image" on the IRISPen control on the Windows taskbar.

Status Window

Refer to the command "Status Window" on the IRISPen control on the Windows taskbar.

Save Settings on Exit

This option guarantees that the current settings are saved as default settings when the user quits the IRISPen software. These settings will be loaded into memory at the next startup.

The default settings are saved in the file DEFAULT.IPC. When this file is erased, the factory settings apply.

Settings files contain all operational parameters: the reading mode, the document language, the target language for the translation (IRISPen Translator), whether dot matrix reading mode (IRISPen Executive) and the mixed reading mode with secondary bar code reading is enabled (IRISPen Executive), the output target, whether the multiline reading mode is enabled (IRISPen Executive and IRISPen Translator), the programming of pen button clicks and hotkeys, table frame and character conversions, the output subset, alignment options (including starting and ending codes), the bar code and handprinting settings (IRISPen Executve), whether the speech synthesis (IRISPen Executive and IRISPen Translator) or system sound is enabled, whether the clipboard translation (IRISPen Translator), the left-handed scanning and the image inversion are enabled.

Related Command

The command "Open Settings" from the "File" menu loads specific IRISPen configurations.

REGISTER MENU

The "Register" menu allows the user to register his software licence.

Registration Wizard

Starts the registration wizard.

The registration wizard guides the user through the licence registration process comfortably. The user can register in many ways: by sending in the registration card or faxing its electronic counterpart, by calling I.R.I.S. during working hours and by filling out a registration form on the I.R.I.S. web site.

Registering the IRISPen licence will allow I.R.I.S. to keep the user informed of future product developments and related I.R.I.S. products. The registration benefits, including free product support and special offers, are strictly limited to registered users.

Upgrade

Gives some information about the upgrading possibilities. IRISPen users can acquire software-only upgrades to migrate to the multifunctional reading pen

IRISPen Executive and to the translation and language learning tool IRISPen Translator. They can also install some IRISPen options.

The IRISPen product range uses a modular approach: by installing an up-grade, the user adds whatever extra feature he doesn't have yet. He can up-grade from the IRISPen Standard to the IRISPen Executive and to the IRISPen Translator, from the IRISPen Translator to the IRISPen Executive, from an IRISPen Executive to an IRISPen Translator etc.
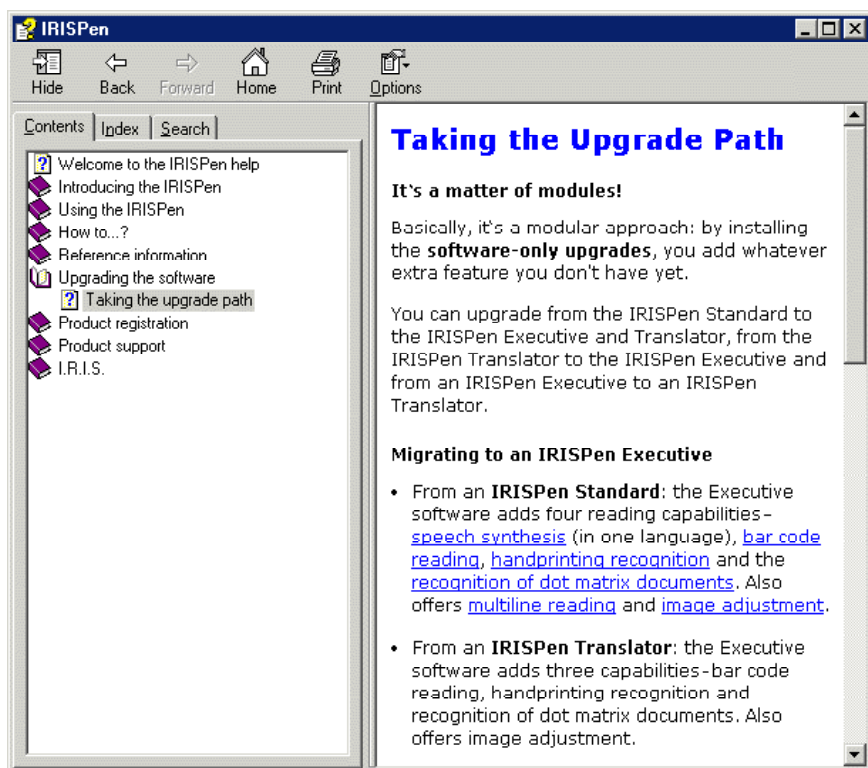
The software-only upgrade to the **IRISPen Executive** offers numerous major add-on capabilities: speech synthesis, bar code reading, handprinting recognition, recognition of dot matrix documents, advanced image adjustment features and multiline scanning. As a result, the IRISPen becomes unquestionably the most versatile scanner available!

The IRISPen Executive offers *speech synthesis* of all read data. The speech synthesis is limited to a single language of the user's choice, but speech modules for additional languages can further extend the user's system. (Ten languages are currently available: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.)

Bar code reading opens a world of new applications to IRISPen owners as their pen scanner becomes an ergonomic and compact bar code scanner for the office environment. Information desks, marketing surveys, customer identification, order picking and many other examples of high accuracy data entry in the office environment can benefit highly from a bar code reading solution.

All widespread bar code standards are supported: Code 39, Code 93, Codabar, EAN-8, EAN-13, UPC and Interleaved 2 of 5. (Support of the Code 128 symbology is optional.) The user can, but does not have to specify the symbology: the user-friendliness is such that the IRISPen automatically detects which bar code standard has been scanned. There is no programming or parameter setting involved in its operation, its use is automatic and transparent.

The user can even enable a mixed reading mode where bar codes are read alongside other data - for instance "normal" text. The IRISPen "auto-detects" which type of data is being read. The user can for instance scan address labels where the printed data is complemented by a bar code without ever changing the reading mode!

Encoding forms and labels can also benefit greatly from numeric *handprinting recognition*. This proprietary recognition technology is based on more than 1 million writing samples! All natural writing styles - American and European - are supported; contrary to most other handprinting recognition solutions, no training phase or imposed style is requied.

The IRISPen Executive also offers recognition of *dot matrix* printed documents. Dot matrix symbols are made up of isolated, separate dots, and highly specialized recognition routines are required to recognize them.

Multiline scanning increases the "user-friendliness" of the IRISPen: it is most useful when the user reads text excerpts. Several scans get "buffered" before the reading result is inserted: the user presses the pen button to send entire paragraphs consisting of several lines to his Windows applications at once. The multiline reading mode also ensures a natural prosody: when the text-to-speech module has to pronounce individual lines, rather than full sentences, the intonation may be somewhat less than "human".

The IRISPen Executive is equipped with advanced *imaging features* to adjust the scanning to the color tones of the documents to be scanned. As the user optimizes the images generated by the pen scanner, a much larger range of documents becomes legible. Think for example of documents where the colored background would otherwise render the recognition impossible.

The software-only upgrade to the **IRISPen Translator** offers the following major add-on capabilities: speech synthesis in two languages, bidirectional sentence translation of a language pair and multiline scanning.

The IRISPen Translator was designed for language learning and translation aid: the user encodes and translates foreign language texts by scanning them, and he hears the words spoken.

The IRISPen Translator provides bidirectional, real-time *sentence translation* of all scanned text. The translation is displayed in a "pop-up" translation window.

The IRISPen Translator software also offers *clipboard translation*: the user translates texts without scanning, any text to be translated is simply copied to the clipboard. Think of texts entered on the keyboard, wordprocessor files, e-mail messages, Internet web pages etc.

The IRISPen Translator English-French translates English into French and vice versa. Extra translation libraries can be installed. The currently available versions are: English-French, English-German, English-Italian, English-Portuguese

and English-Spanish. Contact I.R.I.S. to learn the availability of new language pairs.

The speech synthesis opens the door to a world of new applications. To learn how to speak a language, the user can just scan a text and hear the PC pronounce the words. The speech feature is offered for the language pair: the English-French version of the IRISPen Translator for instance has the English and French speech libraries. (Speech modules for additional languages can further extend the user's system. Ten languages are currently available: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.)

The *multiline scanning* is mandatory for the IRISPen Translator. As sentences in most cases comprise several lines of text, "multiline" scanning allows to buffer several lines before the output process - text insertion, pronunciation and translation - are triggered. Otherwise the system would pronounce partial sentences and (try to) translate ungrammatical text lines as the user scanned them, with unavoidably weak results.

Finally, there are also some **options** available for the IRISPen product range.

The *speech synthesis* of the IRISPen Executive is limited to a single language chosen by the user, the IRISPen Translator is equipped with the speech libraries of the two languages in the language pair. The user can extend his system effortlessly by installing text-to-speech modules for additional languages. Ten languages are currently available: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.

Secondly, the IRISPen Executive supports all widespread bar code symbologies, but support of *Code 128*, a specialized bar code used in the financial world, is optional.

The same goes for the recognition of the specialized *banking fonts* (OCR-A, OCR-B, E13B and CMC-7).

HELP MENU

The "Help" menu gives the user on-line information about the IRISPen and its use. Some last-minute information may not be included in this manual. The user is therefore recommended to check the on-line help system for additional information on the IRISPen's capabilities.

<u>Help Topics</u>

Allows to navigate through the on-line help system.

The available information cannot only be consulted on-screen but also printed out.

USER'S GUIDE
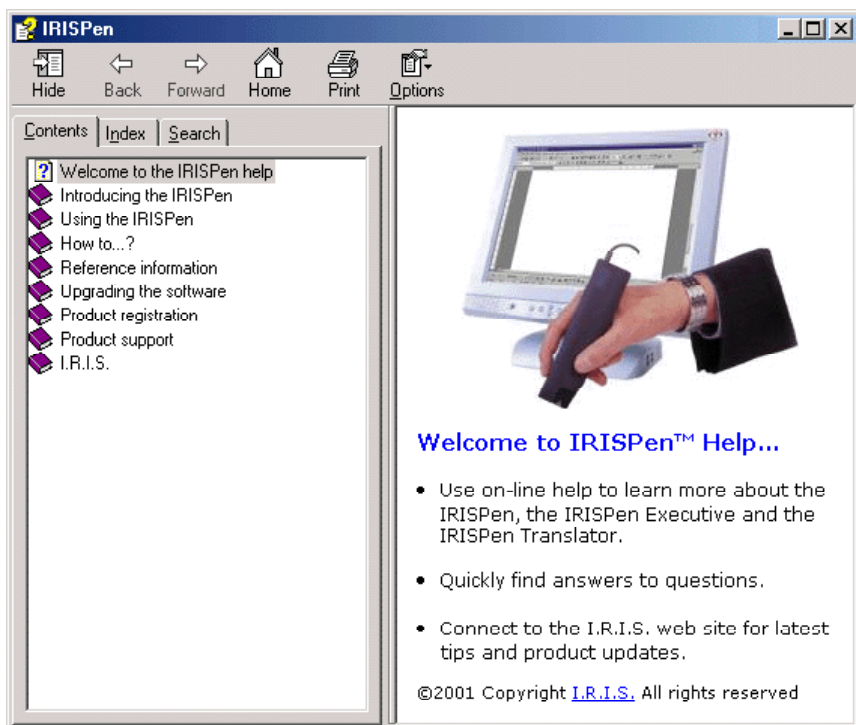
Related Command

The user can also access the on-line help by right-clicking the IRISPen control on the Windows taskbar and selecting the command "Help Topics".

Product Support

Informs the user how to get IRISPen product support.
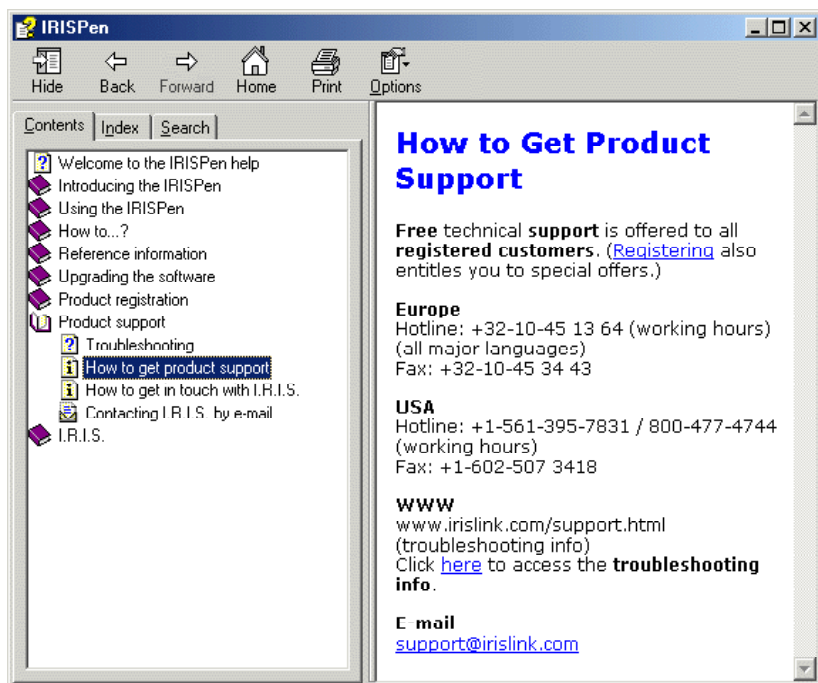
Contact I.R.I.S.

Gives the coordinates of the I.R.I.S. offices.
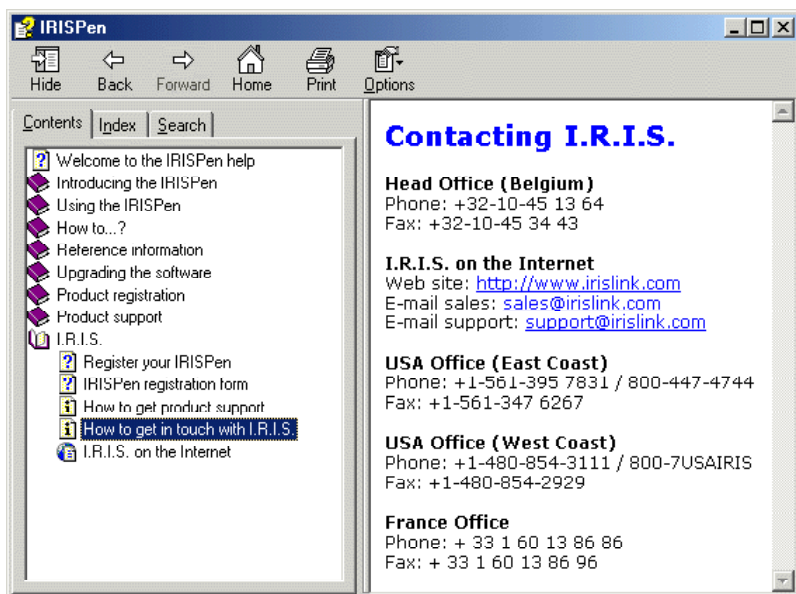
I.R.I.S. on the Internet

Gives access to the I.R.I.S. home page (www.irislink.com).

The user can also access the application icon "I.R.I.S. on the Internet" in the submenu "I.R.I.S. Applications - IRISPen" under the "Programs" menu to get direct access to the I.R.I.S. home page.

About the IRISPen

Opens the IRISPen title screen, where the current version and the copyrights are mentioned.

The user can get direct access to the I.R.I.S. home page by clicking on the URL (www.irislink.com).

Appendix A
PRODUCT DESCRIPTION

This appendix gives a detailed product description of the IRISPen, the IRISPen Executive and the IRISPen Translator.

IRISPEN PRODUCT RANGE

Standard version	pen-sized OCR scanner
	"basic" version
Executive version	multifunctional reading pen
	a software-only upgrade from the IRISPen Standard and from the IRISPen Translator is available
Translator version	translation and language learning tool
	a software-only upgrade from the IRISPen Standard and from the IRISPen Executive is available
Banking version	contact I.R.I.S. to get more information
IRISPen options	speech synthesis libraries for extra languages
	bar code reading of Code 128 symbology

SYSTEM REQUIREMENTS

IRISPen Standard
❑ a 486 based Intel PC or compatible. A Pentium based PC is recommended
❑ 50 MB hard disk space. 40 MB of disk space suffices when the electronic manual is left on the CD-ROM
❑ the Windows ME, 2000 or 98 operating system
❑ a free USB port to connect the IRISPen scanner
❑ a 256 color monitor to display the greyscale images generated by the greyscale pen scanner

IRISPen Executive
❑ a 486 based Intel PC or compatible. A Pentium based PC is recommended
❑ 60 MB hard disk space. 50 MB of disk space suffices when the electronic manual is left on the CD-ROM
❑ the Windows ME, 2000 or 98 operating system
❑ a free USB port to connect the IRISPen scanner
❑ the speech synthesis feature requires that your PC is equipped with a Windows compatible sound card. For quality sound, preference is given to a 64-bit sound card

IRISPen Translator
❑ a 486 based Intel PC or compatible. A Pentium based PC is recommended
❑ 80 MB hard disk space. 70 MB of disk space suffices when the electronic manual is left on the CD-ROM
❑ the Windows ME, 2000 or 98 operating system
❑ a free USB port to connect the IRISPen scanner
❑ the speech synthesis feature requires that your PC is equipped with a Windows compatible sound card. For quality sound, preference is given to a 64-bit sound card

IRISPen options
❑ to install additional speech modules, 10 MB of free hard disk space is required per language
❑ 10 languages are currently available: English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish

Mac OS version
❑ version for Mac OS platform available
❑ cross-platform compatibility of IRISPen hardware with all Macintosh desktop computers

IRISPEN STANDARD

Font support	proprietary omnifont PCR-II ("Pen Character Recognition") technology
	self-learning technique based on contextual analysis
Linguistic support	55 languages (American English, British English, Afrikaans, Albanian, Basque, Brazilian, Breton, Bulgarian, Byelorussian, Catalan, Corsican, Croatian, Czech, Danish, Dutch, Estonian, Finnish, French, Frisian, German, Greek, Hungarian, Icelandic, Indonesian, Irish (Gaelic), Italian, Lithuanian, Luxemburgh, Macedonian, Malay, Mexican, Norwegian, Nynorsk, Polish, Portuguese, Romanian, Russian, Scottish (Gaelic), Serbian, Slovak, Slovenian, Spanish, Swahili, Swedish, Tagalog, Turkish and Ukranian)
	7 mixed language modes (Bulgarian-English, Byelorussian-English, Greek-English, Macedonian-English, Russian-English, Serbian-English and Ukranian-English)

linguistic modules using knowledge about syllables and a 150,000 word lexicon per language

Character modes alphanumeric and numeric
customized character sets
character conversion
table frame detection

Character sizes 8 to 22 point (0.11 to 0.31" or 0.28 to 0.78 cm.)

Graphics greyscale and black-and-white Windows bitmaps (BMP) and proprietary IRISPen graphics
copying of graphics to clipboard for direct pasting and saving of graphics to file

Scanning features greyscale scanner
reads low-contrast images, wavy lines and distorted characters
lineskew tolerance 10%
word slope tolerance 20%
central line tracking
horizontal scanning resolution 300 dpi
vertical scanning resolution 320 dpi
left-to-right and right-to-left scanning

Scanning speed up to 3.15" (8 cm.) per second
speed detection by scanner's roller
speed monitoring by software

USER'S GUIDE

Processing speed   up to 1,000 characters per second

Line length   up to 12" (30.5 cm.) per scan

Color tones   greyscale scanning copes with low-contrast images

uses intelligent, automatic image processing techniques

advanced manual brightness adjustment and image despeckling (IRISPen Executive)

red and yellow dropout color

image inversion of light text on a dark background

Connection   "hot pluggable" device connects to the PC's USB port

powered by the PC via the USB port (no need for an external power supply)

Cable length   180 cm.

USB devices can be reconnected using a USB hub

Lifespan   3.75 miles (6 km.)

Linguistic versions   English, German, French, Dutch, Italian and Spanish

contact I.R.I.S. to learn the availability of additional versions

IRISPEN EXECUTIVE

Executive version   multifunctional reading pen

adds speech synthesis, bar code reading, numeric handprinting recognition, recognition of dot matrix documents and mutliline scanning

a software-only upgrade from the IRISPen Standard and IRISPen Translator is available

Speech synthesis   clear, natural sounding auditive feedback of recognized data

"hands- and eyes busy" data collection and multimedia applications

high-quality technology from Lernout & Hauspie

requires Windows compatible sound card

a single version of the speech library is supplied, the language depends on the user's choice

extension with other languages through add-on modules

multilingual support of 10 languages (English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish)

Bar code reading   fast and extremely accurate data collection

support of all common bar code symbologies (Code 39, Code 93, Interleaved 2 of 5, Codabar, UPC, EAN-8 and EAN-13)

support of Code 128 symbology is optional

user control and automatic detection of bar code symbology

skew- and distortion-tolerant proprietary PBR ("Pen Bar Code Reading") technology

Numeric handprinting recognition

accurate, flexible recognition of numerals and the symbols , (comma), . (dot) and - (hyphen) in all natural American and European writing styles

financial applications and form and label reading

immediate recognition without training phase or imposed writing style

style-independent recognition technology based on proprietary neural network derived from over 1 million writing samples in American and European writing styles

IRISPEN TRANSLATOR

Translator version

translation and language learning tool

adds speech synthesis, sentence translation in real-time and multiline scanning

a software-only upgrade from the IRISPen Standard and IRISPen Executive is available

Speech synthesis

clear, natural sounding pronunciation of foreign languages

high-quality, native speaker-based technology from Lernout & Hauspie

requires a Windows compatible sound card

two linguistic versions of the speech library are supplied, the languages depend on the language pair of the IRISPen Translator

extension with other languages through add-on modules

multilingual support of 10 languages (English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish)

Translation   bidirectional sentence translation of language pair

includes grammatical knowledge (tense, person, gender etc.)

fully automatic, real-time translation without any keyboard input

translation of all text sent to clipboard (text entered on keyboard, text files, e-mail messages, Internet web pages etc.)

first draft, "information only" translation of businesslike, non-emotive sentences

Linguistics   extensive use of linguistic intelligence at text recognition, speech synthesis and translation stage

text recognition enhanced by linguistic databases containing 150,000 words

word-based, contextual analysis complemented by syntactical rules

speech synthesis uses native speaker-based speech technology from Lernout & Hauspie

context-sensitive solving of ambiguous forms ("homographs")

incorporates knowledge of abbreviations and normalized text strings (telephone number, bank account and social security numbers, dates, time indications and currencies)

translation: bidirectional sentence translation

incorporates knowledge of phrases, abbreviations and normalized text strings

Translation speed   up to 50 words per second

Language pairs   English-French, English-German, English-Italian, English-Portuguese and English-Spanish

contact I.R.I.S. to learn the availability of new language pairs

IRISPEN FOR BANKING WORLD

Reading of financial documents

high-accuracy reading of specialized banking fonts (OCR-A, OCR-B, E13B and CMC-7)

support of bar code symbology Code 128

contact I.R.I.S. to get more information

IRISPEN OPTIONS

Speech synthesis   IRISPen Executive: a single version of the speech library is supplied, the language depends on the user's choice

IRISPen Translator: the speech libraries of the language pair are supplied

extension with other languages through add-on modules

multilingual support of 10 languages (English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish)

to install additional speech modules, 10 MB of free hard disk space is required per language

Bar code reading   support of bar code symbology Code 128

Appendix B
SYNTAX RULES
SPEECH SYNTHESIS

I.R.I.S. test sheets are provided with the IRISPen Executive and the IRISPen Translator to get a better idea of the highly advanced nature of "Read & Speak".

They allow you to experiment with the text-to-speech capability of your IRISPen and get aware of the linguistic rules you should respect when you want to correctly read "normalized" character strings such as telephone numbers, bank account numbers, date and time indications etc.

To ensure that the speech feature will produce correct results for such character strings, you should respect these rules when your documents - forms, labels etc. - are designed and produced. In other words, these syntax rules tell you how English dates, French telephone numbers and German bank account numbers should be composed to be recognized as such.

These rules are discussed below per language - it is obvious that such writing rules differ per country. One general rule is this: insert periods between the letters of a word to have the word read letter-per-letter. "N.A.T.O." for example yields a different result than "NATO"!

The speech synthesis of the IRISPen Executive is limited to a single language of your choice, the IRISPen Translator is equipped with the speech libraries of the two languages in the language pair. You can extend your system effortlessly by installing text-to-speech modules for additional languages. Ten languages are currently available: (American) English and British English, Brazilian, Dutch, French, German, Italian, Mexican, Russian and Spanish.

AMERICAN ENGLISH SYNTAX RULES

Numbers

Cardinal numbers of up to 15 digits are pronounced as full numbers, longeer digit strings are pronounced digit by digit. Numbers that start with a zero are equally pronounced digit by digit.

123456789012345	One hundred twenty-three trillion four hundred...
1234567890123456	One two three four five six seven eight nine zero...
007	Zero zero seven

Comma's may be used to separate groups of three digits.

1,234	One thousand two hundred thirty-four
123,456,789,012,345	One hundred twenty-three trillion four hundred...

Numbers ranging from 1,100 to 9,999 are treated as figures when a comma is inserted, and as **years** when no comma is present.

1,995	One thousand nine hundred ninety-five
1995	Nineteen ninety-five
10,999	Ten thousand nine hundred ninety-nine

Decimal numbers may contain up to 15 digits before the decimal point. Comma's may be used to separate groups of three digits (before the decimal point). When more than two digits occur after the decimal point, these are pronounced digit by digit.

104.2	One hundred four point two
.02	Point zero two
.001	Zero point zero zero one

The strings 1st, 2d, 2nd, 3d, 3rd and all cardinal numbers of up to 15 digits followed by an ordinal suffix are pronounced as **ordinal numbers.**

1st, 2d, 2nd and 3rd	First, second, second and third
22nd and 25th	Twenty-second and twenty-fifth

4,567th	Four thousand five hundred sixty seventh

Roman numbers are also pronounced as ordinal numbers. (The symbols I, V and X in isolation are *not* identified as ordinal numbers.)

XX	Twentieth
Queen Mary I	Queen Mary the first

Digit strings composed of up to 3 digits, a slash and up to 5 digits are pronounced as **fractions**.

1/3 to 1/2	One third to one half
3/2	Three halves
123/45678	One hundred and twenty-three over forty-five thousand six hundred seventy eight

Telephone Numbers

Telephone numbers are read digit by digit. A space, a period, a slash or a hyphen is used to separate digit groups. The first digit group can be surrounded by parentheses and will be preceded by the words "area code".

(617) 932 9209	Area code six one seven nine three two nine two...
(564) 123.4567	Area code five six four one two three four five six...
65/234/5678	Area code six five two three four five six seven eight
75-234-5678	Area code seven five two three four five six seven...

If the first digit group is 800 or 900, it is pronounced as one number.

800 123 4567	Eight hundred one two three four five six seven

International telephone formats are recognized too. The first digit group is preceded by the words "country code" and pronounced as one number. The rest of the telephone number is read digit by digit.

+49 (228) 550-381	Country code forty-nine area code two two eight...
+30 61 997 283	Country code thirty area code six one nine nine...
(33) 183 447-593	Country code thirty-three area code one eight...

Bank Account and Social Security Numbers

Bank account numbers and social security numbers are pronounced digit by digit when hyphens separate the digit groups.

13456-000-123476   000-1234564-87   049-65-5432

Dates

Dates are recognized in the **American format** - where the month precedes the day - and the **European format** - where the day precedes the month. The day and the month are composed of one or two digits, the year of two or four digits. The digit groups are separated by a slash or a hyphen.

American format

12/25/99	December twenty-fifth ninety-nine
02/06/00	February six two thousand
4-17-2000	April seventeenth nineteen two thousand

European format

13/12/99	December thirteenth ninety-nine
25/12/00	December twenty-fifth two thousand

Dates can also be specified in "written", **alphabetic formats**.

Thu, 21 October 99	Thursday the twenty-first of October ninety-nine
Fri. 13 Feb 2000	Friday the thirteenth of February two thousand
Mon, Oct 27 1998	Monday, October twenty-seventh nineteen ninety-eight
December 27, 2000	December twenty-seventh two thousand
27th Jan. 2001	The twenty-seventh of January two thousand one
10th May 99	The tenth of May ninety-nine
July 4, 00	July fourth two thousand
7 April	The seventh of April
02 May	The second of May
Oct 22	October twenty-second

The abbreviations for the **decennia** are recognized.

60s	70s	80s	90s

Time Indications

Time indications are recognized when one of the following formats is respected.

Hours only

9AM	12 pm

Hours and minutes

9:00	Nine o'clock
11:15	Eleven fifteen
from 5:20 am to 8:30 pm	From five twenty am to eight thirty pm

Hours, minutes and seconds

Durations are indicated by adding a third group of digits. Two formats are supported: the digit groups are separated by ":" or you get the format *hour:minutes'seconds''*.

12:15:20	Twelve hours, fifteen minutes and twenty seconds
11:55'46"	Ten hours, fifty-five minutes and forty-six seconds

Currencies

Currency indications with $, USD. dol., DM, £, ¥ and Schil. are recognized.

The $, £ and ¥ symbol may precede and follow the numeral (without a space).

$50 and 30$	Fifty dollars and thirty dollars
£5 and 7£	Five pounds and seven pounds
¥1 and 2¥	One yen and two yen

Have the currency precede the numeral to specify decimal amounts. Decimal numbers up to 12 digits and 2 digits are the decimal point will be pronounced correctly.

$1.25	One dollar and twenty-five cents

| £3,100.50 | Three thousand one hundred pounds and fifty pence |
| ¥2.15 | Two yen and fifteen sen |

The currency abbreviations DM, USD, dol. and Schil. must follow the numeral with a space added.

10 USD	Ten US dollars
100 dol.	One hundred dollars
25 DM	Twenty-five Deutsche marks
88 Schil.	Eighty-eight Schillings

Other currencies must be written in full and follow the numeral. Alternatively, you can program the IRISPen to for instance replace "F" by "francs" with the "Character Conversion" command under the "Settings" menu.

| 20 francs | 100 liras |

Abbreviations

A standard dictionary includes all common abbreviations. Ambiguous abbreviations are solved based on the linguistic context: "Dr." is pronounced as "doctor" when followed by a capital letter and as "drive" in all other cases, "St." is pronounced as "saint" when followed by a capital letter and as "street" in other cases.

Dictionary contained abbreviations

ASAP	TTS	Dr.	e.g.	etc.	Ltd.
Co.	Corp.	Bros.	Jr.	Sgt.	Lt.
Col.	Gen.	Mr.	Mrs.	Ms.	ca.
St.	mm.	cm.	Tel.	Inc.	excl.
gr	kg	i.e.	Dept.	Hr.	Mgr.
Av.	Ave.	c/o	Pres.	NY	N.Y.
L.A.	Br.	WWW			

The abbreviations for months are recognized.

| Jan. | Aug. | Oct. | Dec. |

Context rules

Dr. Johnson lives in Elm Dr.	Doctor Johnson lives in Elm drive
St. Peter's St.	Saint Peter's street
Dr. Johnson lives in St. Peter's St.	Doctor Johnson lives in Saint Peter's street

Acronyms

Acronyms are pronounced as normal words.

NASA	NATO	IRIS	UNESCO

Initialisms

Initialisms are pronounced letter by letter.

FBI	CIA	IRS	FDA	EPA	NCO
CDC	TV	Ph.D.	MBA	USSR	GMT
GQ	ABC	BC	BSc.	CEO	PC
PR	LP	CD	MC	OK	VP
FCC	PD	VIP	GOP		

BRAZILIAN SYNTAX RULES

Numbers

Cardinal numbers of up to 15 digits are pronounced as full numbers, longer digit strings are pronounced digit by digit. Leading zero's are equally pronounced digit by digit.

123456789012345	1234567890123456	007

Periods may be used to separate groups of three digits.

1.234	123.456.789.012.345

Decimal numbers may contain up to 15 digits before and after the comma. Periods may be used to separate groups of three digits (before the comma).

When more than two digits occur after the decimal point, these are pronounced digit by digit.

104.2 5.25 5.151

All cardinal numbers from 1 to 999 followed by ª or º are pronounced as **ordinal numbers**. So are all cardinal numbers of up to 15 digits that start with a digit other than zero and are followed by ª or º. Periods may be used to separate groups of three digits.

2ª 134º 1,000

Depending on the context, the **Roman numbers** will be pronounced as ordinal numbers in male and female form or as cardinal numbers.

Rainha Elisabete II Rainha Elissabete Segunda

Capítulo XX Capítulo vigésimo

Século XXI Século vinte e um

Telephone Numbers

The following formats are supported for telephone numbers: telephone numbers with an optional operator code and a country code, telephone numbers with an optional operator code and an area code, telephone numbers preceded by the abbreviation "t", "tel", "telef", "fone" and "fax".

The operator code, the country code and the area code may be surrounded by parentheses; they may also be separated using a hyphen or a space. A hyphen or a space can also separate the last four digits of the telephone number. Don't change the type of separator within a telephone number!

tel. 6941 2786 fone 550-2812 011 6942 5781

(015) (32) 950-2742 0015-1-212-555-7781

Dates

Dates are recognized in a number of formats. The day precedes the month. The day and the month are composed of one or two digits, the year of two or four digits. When the year is indicated by two digits, it may be preceded by a quote. The digit groups are separated by a slash or a hyphen.

08-02-'99 16/05/2000

Dates can also be specified in "written", **alphabetic formats**.

8 jul. 98   30 abril 2000

Time Indications

Time indications are recognized when one of the following formats is respected.

Hours only

One or two digits are used for the hour which is followed by "h" or "hs" (without a space).

06h   12hs

Hours and minutes

One or two digits are used for the hour followed by "h" (without a space); two digits are used for the minutes followed by "min" (without a space).

6h32min   23hl0min

One or two digits are used for the hour followed by a colon or period; two digits are used for the minutes followed by "h", "hs", "min", "am" or "pm".

2:30 min   10.30am   12:00hs

One or two digits are used for the hour followed by a colon; two digits are used for the minutes.

23:55   8:30

Hours, minutes and seconds

Durations are indicated by adding a third group of digits. Use two digits for the seconds and use a colon to separate the digit groups.

21:00:32   9:23:43

Currencies

A wide range of currencies is recognized: R$, Cr$, NCr$, Cz$, US$, $ G/, Ur$, Euro, Eur and ¥. The currencty symbol must precede the numeral (without a space).

US$40   R$30,25   Cr$3.000,00

Currency indicators can be followed by one to six digits (with one or two digits after the comma) and the number indicator "milhão", "bilhão", "trilhão", "mi", "bi" or "tri".

R$23 milhões   US$10,5 bi

Other currencies must be written in full and follow the numeral. Alternatively, you can program the IRISPen to for instance replace "F" by "francos" with the "Character Conversion" command under the "Settings" menu.

20 francos   100 liras

Abbreviations

A standard dictionary includes all common abbreviations. Ambiguous abbreviations are solved based on the linguistic context: "km" is pronounced as "quilômetro" when it is preceded by the number 1, but pronounced as "quilômetros" when preceded by another number. Similarly, "cap." is pronounced "capítulo" when followed by a number, but "capitão" in alI other contexts.

Dictionary contained abbreviations

sr.   dr.   obs.   ex.   etc.

jr.   km   cap.

Context rules

21 km   Vinte e um quilômetros

1 km   Um quilômetro

cap.4   Capítulo quatro

cap. Souza Arantes   Capitão Souza Arantes

Acronyms

Acronyms are pronounced as normal words.

IRIS   UNESCO   USP   UNICEF

Initialisms

Initialisms are pronounced letter by letter.

HIV   PT

BRITISH ENGLISH SYNTAX RULES

Numbers

Cardinal numbers of up to 15 digits are pronounced as full numbers, digit strings that contain more than 15 digits are pronounced digit by digit. Numbers that start with a zero are equally pronounced digit by digit.

123456789012345	One hundred twenty-three trillion four hundred...
1234567890123456	One two three four five six seven eight nine zero one...
007	Zero zero seven

Comma's may be used to separate groups of three digits.

1,234	One thousand two hundred thirty-four
123,456,789,012,345	One hundred twenty-three trillion four hundred...

Numbers ranging from 1,100 to 9,999 are treated as figures when a comma is inserted, and as **years** when no comma is present.

1,995	One thousand nine and hundred ninety-five
1995	Nineteen ninety-five
10,999	ten thousand nine hundred ninety-nine

Decimal numbers may contain up to 15 digits before the decimal point. Comma's may be used to separate groups of three digits (before the decimal point). When more than two digits occur after the decimal point, these are pronounced digit by digit.

104.2	One hundred and four point two
.02	Zero point zero two
.001	Zero point zero zero one

The strings 1st, 2d, 2nd, 3d, 3rd and all cardinal numbers of up to 15 digits followed by an ordinal suffix are pronounced as **ordinal numbers**.

1st, 2d, 2nd and 3rd	First, second, second and third
22nd and 25th	Twenty-second and twenty-fifth

4,567th	Four thousand five hundred and sixty seventh

Roman numbers are also pronounced as ordinal numbers. (The symbols I, V and X in isolation are *not* identified as ordinal numbers.)

XX	Twentieth

Digit strings composed of up to 3 digits, a slash and up to 5 digits are pronounced as **fractions**.

1/3 to 1/2	One over three to one half
3/2	Three halves
123/45678	One hundred and twenty-three over forty-five thousand six hundred and seventy eight

Telephone Numbers

Telephone numbers are read digit by digit. A space, a period, a slash or a hyphen is used to separate digit groups. The first digit group can be surrounded by parentheses and will be preceded by the words "area code".

(617) 932 9209	Area code six one seven nine three two nine two zero...
(564) 123.4567	Area code five six four one two three four five six...
65/234/5678	Area code six five two three four five six seven eight
75-234-5678	Area code seven five two three four five six seven...

If the first digit group is 800 or 900, it is pronounced as one number.

800 123 4567	Eight hundred one two three four five six seven

International telephone formats are recognized too. The first digit group is preceded by the words "country code" and pronounced as one number. The rest of the telephone number is read digit by digit.

+49 (228) 550-381	Country code forty-nine area code double two eight...
+30 61 997 283	Country code thirty area code six one double nine...
(33) 183 447-593	Country code thirty-three area code one eight three...

Bank Account Numbers

Bank account numbers are pronounced digit by digit when hyphens separate the digit groups.

000-1234564-87   049-65-5432

Dates

Dates are recognized in the **European format**: the day precedes the month. The day and the month are composed of one or two digits, the year of two or four digits. The digit groups are separated by a slash or a hyphen.

13/12/99	The thirteenth of December ninety-nine
25/12/00	The twenty-fifth of December two thousand

Dates can also be specified in "written", **alphabetic formats**.

Fri. 13 Feb 2000	Friday the thirteenth of February two thousand
Mon, Oct 27 1998	Monday, October twenty-seventh nineteen ninety-eight
December 27, 2000	Twenty-seventh two thousand
27th Jan. 2001	Twenty-seventh of January two thousand one
10th May 99	The tenth of May ninety-nine
July 4, 2000	Fourth July two thousand
7 April	The seventh of April
02 May	The second of May
Oct 22	October twenty-second

The abbreviations for the **decennia** are recognized.

60s   70s   90s

Time Indications

Time indications are recognized when one of the following formats is respected.

Hours only

9AM   12 pm

Hours and minutes

10:30	Ten thirty
11:15	Eleven fifteen
from 5:20 am to 8:30 pm	From five twenty am to eight thirty pm

Hours, minutes and seconds

Durations are indicated by adding a third group of digits. Two formats are supported: the digit groups are separated by ":" or you get the format *hour:minutes'seconds''*.

| 12:15:20 | Twelve hours, fifteen minutes and twenty seconds |
| 11:55'46" | Eleven hours, fifty-five minutes and forty-six seconds |

Currencies

Currency indications with $, £ and ¥ are recognized. The $, £ and ¥ symbol may precede and follow the numeral (without a space).

$50 and 30$	Fifty dollars and thirty dollars
£5 and 7£	Five pounds and seven pounds
¥1 and 2¥	One yen and two yen

Have the currency precede the numeral to specify decimal amounts. Decimal numbers up to 12 digits and 2 digits are the decimal point will be pronounced correctly.

$1.25	One dollar and twenty-five cents
£3,100.50	Three thousand one hundred pounds and fifty pence
¥2.15	Two yen and fifteen sen

Other currencies must be written in full and follow the numeral. Alternatively, you can program the IRISPen to for instance replace "F" by "francs" with the "Character Conversion" command under the "Settings" menu.

20 francs   100 liras

Abbreviations

A standard dictionary includes all common abbreviations. Ambiguous abbreviations are solved based on the linguistic context: "Dr." is pronounced as "doctor" when followed by a capital letter and as "drive" in all other cases, "St." is pronounced as "saint" when followed by a capital letter and as "street" in other cases.

Dictionary contained abbreviations

ASAP	TTS	Dr.	e.g.	etc.	Ltd.
Bros.	Jr.	Sgt.	Lt.	Col.	Gen.
Mr.	Mrs.	ca.	St.	Eng.	mm.
cm.	Tel.	Inc.	excl.	gr	kg
i.e.	Dept.	Hr.	Av.	Ave.	Pres.

The abbreviations for months are recognized.

Jan.	Aug.	Oct.	Dec.

Context rules

Dr. Johnson lives in Elm Dr.	Doctor Johnson lives in Elm drive
St. Peter's St.	Saint Peter's street
Dr. Johnson lives in St. Peter's St.	Doctor Johnson lives in Saint Peter's street

Acronyms

Acronyms are pronounced as normal words.

NASA	IRIS	UNESCO

Initialisms

Initialisms are pronounced letter by letter.

EEC	BBC	VAT	TV	MP	Ph.D.
MBA	USSR	GMT	GQ	BC	BSc.
GB	UK	CEO	PC	PR	LP

CD   MC   OK   VP   VIP

DUTCH SYNTAX RULES

Numbers

Cardinal numbers of up to 15 digits are pronounced as full numbers, longer digit strings are pronounced digit by digit. Periods can be used to separate the digit groups. Leading zeros are pronounced separately.

123456789012345	Honderd drieëntwintig triljoen vier honderd ...
123456	Honderd drieëntwintig duizend vierhonderd zesenvijftig
1234567890123456	Eén twee drie vier vijf zes zeven acht negen nul één...
007	Nul nul zeven

Periods may be used to separate groups of three digits.

123.456	Honderd drieëntwintig duizend vierhonderd zesenvijftig
1234.56.789.012.345	Honderd drieëntwintig triljoen vier honderd

Decimal numbers may contain up to 15 digits before the comma and up to 15 digits after the comma. Periods may be used to separate groups of three digits (before the comma). The digit string before the comma is pronounced as a full number. The digit string after the comma is pronounced digit by digit when it contains more than two digits.

150,65	Honderd vijftig comma vijfenzestig
1,2	Eén comma twee
0,02	Nul comma nul twee
1,234	Een comma twee drie vier

Alphanumeric Strings

With alphanumeric strings, a combination of letters and digits without spaces, the alphabetic part is spelled and the numeric part is read as a full number.

KX4956   125DC20

Telephone Numbers

Telephone numbers are recognized when parentheses, slashes or hyphens are used to separate the country code and/or area code from the actual telephone number. Spaces and periods are allowed to separate groups of digits. The telephone numbers are pronounced in groups of two or three digits, a pause is inserted where a period or space occurs.

030/288 65 27	(030)2886527	030/288.65.27
0031/30/288.65.27	(+32)(52)226677	05780-12401

Telephone numbers may be preceded by the abbreviation "tel.".

Tel. 229 30 11   tel. (057)22.16.07

Bank Account Numbers

Bank account numbers are pronounced digit by digit when hyphens separate the digit groups. The number will be read in groups of two or three digits.

000-1254887-87   123-4528098-34

Dates

Dates are recognized in the following format: day precedes month precedes year. The day and the month are composed of one or two digits, the year of two or four digits. The digit groups are separated by a period, a slash or a hyphen.

1/1/2001	Eén januari twee duizend één
25.12.2000	Vijfentwintig december twee duizend
13-12-99	Dertien december negen en negentig
25/12/00	Vijfentwintig december twee duizend

Year that are composed of two digits, may be preceded by a quote.

01/01/'00	Eén januari twee duizend

Dates can also be specified in "written", **alphabetic formats**.

13 dec. 2000	Dertien december twee duizend
25 febr. 2001	Vijfentwintig februari twee duizend één

Time Indications

Time indications are recognized when one of the following formats is respected. The indicators "u", "u.", "h", "h." and "uur" can be used.

8:15u	10u30		
10:05u.	10u.45		
7 h	10.15h	20:15h	10h50
10:20 PM			

Currencies

All common currency indications are supported. Most currency symbol can precede and follow the numeral.

100 fr.	10 BF.	65 BEF	20 Eur.
25 £	150$	10 DM.	
USD 9	¥5		

Abbreviations

A standard dictionary includes all common abbreviations.

Dictionary contained abbreviations

bv.	b.v.	ir.	ing.	dr.	d.m.v.
m.a.w.	enz.	N.V.	Mr.	Mevr.	blz.
St.	p.a.	resp.	m.i.	e.a.	ca.
z.o.z.	Fr.	incl.	t.o.v.	t.a.v.	t.e.m.
cm.	mm.	excl.	Hr.	Nr	o.l.v.
drs.	lic.	a.s.	e.k.	jl.	min.
max.	vb.				

Acronyms

Acronyms are pronounced as normal words.

NASA	NAVO	HAVO	UNICEF	UNESCO	RIZIV

Initialisms

Initialisms are pronounced letter by letter.

EEG   VDAB

Words consisting of consonants only are considered to be initialisms. So are alphabetic strings (with or without periods) that have two or more uppercase consonants. They are pronounced letter by letter.

TV	BTW	PS	NB	VZW	CV
PR	CD	PC	PK		

FRENCH SYNTAX RULES

Numbers

Cardinal numbers of up to 15 digits are pronounced as full numbers, longer digit strings are pronounced digit by digit. Leading zeros are pronounced separately.

123456789012345   Cent vingt-trois trillions quatre cent cinquante six...

007   Zéro zéro sept

Periods may be used to separate groups of three digits.

1.234   Mille deux cents trente quatre

123.456.789   Cent vingt-trois millions quatre cents cinquante six...

Decimal numbers may contain up to 15 digits before the comma and up to 15 digits after the comma. Periods may be used to separate groups of three digits (before the comma). The digit string after the comma is pronounced as a full number.

100,65   Cent virgule soixante-cinq

1,2   Un virgule deux

Cardinal number followed by the appropriate suffixes are pronounced as **ordinal numbers**.

le 1er et le 10eme   la 21e fois

Roman numbers - numbers up to 38 that consist of combinations of X, V and I - are supported. Roman numbers may be used separately and be combined with proper names.

Louis XIV   XIX

Romans up to 28, followed by "e", "e.", "E" or "E." are read as ordinal numbers.

Xe   Ie.

Telephone Numbers

Telephone numbers are recognized when parentheses, slashes or hyphens are used to separate the country code and/or area code from the actual telephone number. Spaces and periods are allowed to separate groups of digits. The telephone numbers are pronounced in groups of two digits, a pause is inserted where a period, hyphen or space occurs.

04.42.20.84.53   03/88.41.73.00   (01)42 21 84 53

(01) 42 21 84 53   83-50-33-33   83 54 21 73

45 03 37 60   45.03.37.60   45-03-37-60

When telephone numbers are preceded by an abbreviation, the number may no longer be separated by spaces, periods or hyphens.

Tél: 45033760   tél. (03)22712999

The following formats are supported for Belgian telephone numbers - the telephone numbers are pronounced in groups of two or three digits.

Tél. (03)820.02.22   Tél. (03)8200222

Tél. (057) 82.00.22   Tél.02/460.33.97

Bank Account Numbers

Bank account numbers are recognized when hyphens are inserted between the groups of digits. The account numbers are pronounced in groups of two or three digits.

000-1254887-87   123-4528098-34

Dates

Dates are recognized in the following format: day precedes month precedes year. The day and the month are composed of one or two digits, the year of two or four digits. The digit groups are separated by a period, a slash or a hyphen.

1/1/2001	Premier janvier deux mille un
25.12.2000	Vingt-cinq décembre deux mille
13-12-99	Treize décembre quatre vingt dix-neuf
25/12/00	Vingt-cinq décembre deux milles

Dates can also be specified in "written", **alphabetic formats**.

le 1er janv. 99	Le premier décembre quatre vingt dix-neuf
le 25 déc. 2001	Vingt-cinq décembre deux milles un

Time Indications

Time indications are recognized when one of the following formats are respected.

20h30 20 H 30 20.30 h	Vingt heures trente
20.30H 20:30h 20:30 H	
20h30' 20 H 30'	Vingt heures trente minutes
12h00	Midi
00h00	Minuit
2 h	Deux heures
14 H	Quatorze heures
22'	Vingt-deux minutes
1240'	Mille deux cent quarante minutes

Currencies

Currency indications with FF, FRS, FB, fr., \$, £ and ¥ are recognized. The currency symbol must follow the numeral (with a space).

15 FF et 10 FB	50 \$ ou 2¥	2 £ et 2 FRS

Other currencies must be written in full and follow the numeral. Alternatively, you can program the IRISPen to for instance replace "F" by "francs" with the "Character Conversion" command under the "Settings" menu.

Alphanumeric Strings

With alphanumeric strings, a combination of letters and digits without spaces, the alphabetic part is spelled and the numeric part is read as a full number.

KX4956   125DC20

Abbreviations

A standard dictionary includes all common abbreviations.

Dictionary contained abbreviations

M.	mm.	Mme	Mlle	Av.	etc.
BP.	Ste	Mr.	Tél.	s.v.p.	S.A.
Tél	Fr.	Frs	incl.	excl.	cm
kg	Nr.	No.	ex.	min.	max.

Acronyms

Acronyms are pronounced as normal words.

ONU	OTAN	SMIC	EURATOM
UNICEF	UNESCO	FOREM	TEC

Initialisms

Initialisms are pronounced letter by letter. All words consisting of consonants only are considered to be initialisms.

PDG	HLM	TVA	TGV	SNCF	OK
RF	NRF	PS	NB	CE	BCBG
TTC	QC	FF.	FB.	INAMI	PC

GERMAN SYNTAX RULES

Numbers

Cardinal numbers of up to 15 digits are pronounced as full numbers, longer digit strings are pronounced digit by digit. Leading zeros are pronounced separately.

123.456.789.012.345	Hundertdreiundzwanzig Trillionen...
007	Nul nul sieben

Periods may be used to separate groups of three digits.

1.234	123.456.789

Decimal numbers may contain up to 15 digits before the comma and up to 15 digits after the comma. Periods may be used to separate groups of three digits (before the comma). The digit string after the comma is pronounced digit by digit.

100,65	Ein hundert Komma sechs fünf
1,2	Einz Punkt zwei

Cardinal numbers smaller than 32 followed by a period are pronounced as **ordinal numbers**. This does not apply when they are not in the beginning or end of a sentence.

am 15. Dezember 1999

Roman numbers - numbers smaller than 10 that cannot be interpreted as single letters - are supported. Roman numbers may be used separately and be combined with proper names.

Abteilung IV	Hundertwasser II

Telephone Numbers

Telephone numbers are recognized when parentheses, slashes or hyphens are used to separate the country code and/or area code from the actual telephone number. Use spaces and periods to separate groups of digits. The telephone

numbers are pronounced in groups of two or three digits, a pause is inserted where a period, hyphen or space occurs.

(041) 317 11 33	(041) 317.11.33
(03 35) 23 02 16	(03 35) 23.02.16

Bank Account Numbers

Bank account numbers are recognized when hyphens are inserted between the digit groups. A short pause is inserted at the place of a hyphen.

LG Stuttgart 303-504-52	BLZ 900-564-34

Dates

Dates are recognized in the following format: day precedes month precedes year. The day and the month are composed of one or two digits, the year of two or four digits. The digit groups are separated by a period or a slash.

19/12/99	Neunzehnter zwölfter neunundneunzig
9.9.2000	Neunter neunter zweitausend
01.02.97	Erster zweiter siebenundneunzig
am 10.07.99	Am zebten neunundneunzig

Time Indications

Time indications are recognized when one of the following formats is respected.

9:15	09:15 Uhr	4.00 Uhr	22.15 Uhr

Currencies

DM currency indications are recognized when one of the following formats is respected.

DM 40	Vierzig D-Mark
50 DM	Fünfzig D-Mark
DM 10,-	Zehn D-Mark

The Austrian, Swiss and American currencies are also supported.

20 Sfr.	Zwanzig schweizerische Franken
13 ÖS.	Dreizehn Österreichische Schillinge

5 $ Fünf Dollar

Other currencies must be written in full words and follow the numeral. Alternatively, you can program the IRISPen to for example replace "F" by "Franken" with the "Character Conversion" command under the "Settings" menu.

20 Franken   100 Gulden

Abbreviations

A standard dictionary includes all common abbreviations.

Dictionary contained abbreviations

bzw.	derg.	usw.	Fa.	Dr.	ca.
Chr.	KFZ	u.a.	MwSt.	Fr.	inkl.
Str.	mm	cm	Tel.	kg	Nr.
min.	max.				

Some abbreviations are case-sensitive.

PF	Postfach	Pf	Pfennig
Hrsg.	Herausgeber	hrsg.	Herausgegeben
So.	Sonntag		

The abbreviations for months are recognized.

Jan.   Febr.   Dez.

Acronyms

Acronyms are pronounced as normal words.

IRIS

Initialisms

Initialisms are pronounced letter by letter. All words consisting of consonants only are considered to be initialisms. So are abbreviations that contain one or more vowels separated by periods.

NDR	ADAC	LKW	p.A.	CD	ZDF

ITALIAN SYNTAX RULES

Numbers

Cardinal numbers of up to 13 digits are pronounced as full numbers, longer digit strings are pronounced digit by digit. Periods may be used to separate groups of three digits. Leading zeros are pronounced separately.

1234567890123   1.234   123.456.789   007

Decimal numbers may contain up to 13 digits before the comma and up to 13 digits after the comma. Periods may be used to separate groups of three digits (before the comma). The digit string after the comma is pronounced as a full number.

100,50   Cento virgola cinquanta

1,2   Una virgola due

Cardinal number followed by ª or º, "esimo" or "esima" are pronounced as **ordinal numbers**.

1º   14ª   20esimo   100esima

Roman numbers are supported separately and in combination with proper names.

XXII   Piazza XX Settembre   del III capitolo   Carlo V

Telephone Numbers

Telephone numbers are recognized when parentheses or slashes are used to separate the area code from the actual telephone number. Hyphens can be used to separate groups of digits. The digit groups are pronounced separately as full numbers.

052/470489   052/47-04-89   (052) 470489   (052) 47-04-89

Telephone numbers may be preceded by the abbreviations "tel.", "Tel.", "TEL." and "Tel.no.".

Tel. 052/470489   Tel.no. 254/341947

Dates

Dates are recognized in the following format: day precedes month precedes year. The day and the month are composed of one or two digits, the year of two or four digits. The digit groups are separated by a slash or a hyphen.

3/3/99	3/3/00	08/02/99	8/2/2000	01/01/1999
8-2-99	8-2-00	08-02-99	8-2-1999	01-01-1999

The date can be preceded by "l".

l'8-02-99

Time Indications

Time indications are recognized when one of the following formats is respected. The hour indicator contains one or two digits, the minute indicator contains two digits.

8:30	11:15	23:00

The time indication can be preceded by the strings "alle", "delle", "dalle", "h", "le" and "ore.

alle 11:33	ore 11.33	h 11.33

Currencies

Currency indications with L., Lit., LIT., $, £, ¥, DM and Pts. are recognized. The currency symbol L. must precede the numeral (with a space), the currency symbols Lit., LIT. and Pts must follow the numeral.

L.40.000	40.000 Lit.	100.000 LIT.	
$ 40,5	DM 7,50	120 £	¥ 95

Other currencies must be written in full and follow the numeral. Alternatively, you can program the IRISPen to for instance replace "F" by "franco" with the "Character Conversion" command under the "Settings" menu.

Abbreviations

A standard dictionary includes all common abbreviations. Ambiguous abbreviations are solved based on the linguistic context: "km" is pronounced as

"chilometro" when preceded by the number 1 and as "chilometri" in all other cases.

Dictionary contained abbreviations

ecc.	dott.	es.	sig.
USA	p.zza	km	

Acronyms

Acronyms are pronounced as normal words.

RAI   SIP

Initialisms

Initialisms are pronounced letter by letter.

CD	PC	LSD	BMW

MEXICAN SPANISH AND SPANISH SYNTAX RULES

Numbers

Cardinal numbers of up to 8 digits are pronounced as full numbers, longer digit strings are pronounced digit by digit. Leading zeros are pronounced separately.

12.345.678	1.234	123.456	007

Periods may be used to separate groups of three digits, in which case up to 15 digits can be pronounced as full numbers.

1.234   123.456.789.012.345

Decimal numbers may contain up to 8 digits before the comma and up to 8 digits after the comma. When periods are used to separate groups of three digits (before the comma), the decimal number may contain up to 15 digits before the comma and up to 8 digits after the comma.

50,5	Cincuenta coma cinco
1,2	Uno coma dos
12345678,12345678	Doce millones trescientos cuarenta y...

123.456.789.012.345,12345678   Ciento veintitres billones cuatrocientos...

Cardinal numbers followed by the ordinal suffix ° and ᵃ are pronounced as **ordinal numbers**. The numbers 1 and 3 may also be followed by "er".

2°   4ᵃ   3er

Roman numbers are also supported. In isolation, they are pronounced as cardinal numbers.

CXXII   Ciento veintidós

Roman numbers from I (1) to X (10) are pronounced as ordinal numbers when they follow or precede a word that begins with an uppercase letter.

Carlos I	Carlos primero
Isabel II	Isable segunda
Enrique VIII	Enrique octavo
Alfonso XII	Alfonso doce
I Congreso	Primer congreso
IV Muestra	Cuarta muestra

Bank Account Numbers

Bank account numbers are recognized when hyphens are inserted between the digit groups. Groups of up to 8 digits will be pronounced as full numbers, larger groups are pronounced digit by digit.

390-556846570-89   26-847364859-10

Telephone Numbers

Telephone numbers are recognized when parentheses are used to separate the area code from the actual telephone number. Use spaces and hyphens to separate groups of digits. The digit groups are pronounced separately as full numbers.

857-19-24   616-5511

Dates

Dates are recognized in the following format: day precedes month precedes year. The day and the month are composed of one or two digits, the year of two or four digits. The digit groups are separated by a slash or a hyphen.

12/12/93	12-12-99	01-1-00
3-3-95	3/3/95	
8/12/1999	12-12-2000	
01/01/2001	01-01-1995	

Time Indications

Time indications are recognized when one of the following formats is respected. The hour indicator contains one or two digits, the minute indicator contains two digits. The hours can be followed by the abbreviations "h", "H", "h." and "H.": the system then automatically inserts the word "horas" and "minutos" where it is useful.

14:30	14:30 h	14.30 H
14.30 h.	14:30 H.	

Currencies

Currency indications with $, £ and ¥ are recognized. The currency symbols follow the numeral (with a space added).

1 $	Un dólar	100 $	Cien dólares
1£	Un libra	100 £	Cien libras
1 ¥	Un yen	100 ¥	Cien yenes

Other currencies must be written in full and follow the numeral. Alternatively, you can program the IRISPen to for instance replace "F" by "francos" with the "Character Conversion" command under the "Settings" menu.

Abbreviations

A standard dictionary includes all common abbreviations.

Dictionary contained abbreviations

etc.	D.	dcha.	Dr.	ptas.	Tel.

ms	s.	Tel.	mm	cm	gr
kg	Av.	c/	P.D.	PVP	Sr.
Sra.	ej.	ed.	NIF		

Acronyms

Acronyms are pronounced as normal words.

EPTA	NATO	UNICEF	BOE	IVA	RENFE
ONCE	TALGO	INEM			

Initialisms

Initialisms are pronounced letter by letter.

IPC	PC	CD	OK

RUSSIAN SYNTAX RULES

Numbers

Cardinal numbers of up to 15 digits are pronounced as full numbers. Longer digit strings are split into groups of two or three digits - these digit groups are pronounced as full numbers. Leading zeros are pronounced separately.

123456789012345   1234567890123456   007

Periods may be used to separate groups of three digits., in which case up to 15 digits can be pronounced as full numbers.

1.234   123.456.789.012.345

Cardinal numbers are declined in the nominative/accusative case.

Decimal numbers may contain up to 15 digits before the comma and up to 3 digits after the comma. The digits after the comma are pronounced as a fraction in tenths, hundredths or thousandths, as appropriate. If the digit string after the comma contains more than three digits, it is pronounced as a cardinal number.

1,2   100,65

12345678,12345678   123.456.789.012.345,12345678

Decimal numbers are equally declined in the nominative/accusative case.

Bank Account Numbers

Digit strings consisting of 5 to 15 digits are pronounced in clusters of two or three cardinal numbers when they are preceded by one of the following entries.

расчётный счёт	Текущии счёт	
сберегательныll счёт	Сб/счём	
р/с	т/с	сб/с

These entries may be followed by a colon.

расчетный счет: 112-1223445-76

р/с: 2354545

т/с 65008903

You can also use hyphens to separate digit groups.

000-1254887-87   09-12548-1

Telephone Numbers

Telephone numbers are recognized when the country code is preceded by a plus sign (+) and when brackets are used to separate the area code from the actual telephone number. Use hyphens to separate groups of digits.

The digit strings in telephone numbers are read as cardinal numbers. Strings of more than three digits are split into groups of two or three digits. Leading zeros are pronounced separately.

(095)232-92-09   345-66-07

+7 (0822) 46-17-98   +7 (0822) 461 798

Digit strings of five digits or more digits - optionally containing hyphens - are recognized when they are preceded by тел., факс or телефон. These entries may be followed by a colon.

Тел.: 345-66-77   Телефон 46-1798   Факс: 461798

Dates

Dates are recognized in the following format: day precedes month precedes year. The day and the month are composed of one or two digits, the year of two or four digits. The digit groups are separated by a slash or a period. These dates must be followed by the word года or its abbreviation г.

19.12.97 года.	Девятнадцатое декабря девяносто седьмого года
27/04/1998 г.	Двадцать седь.iчое апреля тысяча девятьсот
	Девяносто вось,woro года
5.12.2001 г.	Пятое января две тысячи первого года

Dates can also use the alphabetic format: they are followed by the word года or its abbreviation г.

12 февраля 1998 года.   12 февраля 1998г.   5 января 2001 г.

When used in isolation, dates are declined in the nominative case. In sentence context, dates are pronounced in the genitive/accusative case.

03.04.63 г	3 апреля 1963 года.
Он родился 03.04.63 г.	Он родился 3 апреля 1963 года.

Time Indications

Time indications are recognized when one of the following formats is respected.

Hours, minutes or seconds only

5 час.	Пять часов
1 час	Один час
21 мин.	Двадцать одна минута
16 сек.	Шестнадцать секунд

Hours and minutes

11:15	Одиннадцать пятнадцат
11.15	Одиннадцать пятнадцать
11 час. 15	Мин. одиннадцать часов пятнадцать минут

Hours, minutes and seconds

15:30:45	Пятнадцать тридцать сорок пять
11 час. 15 мин. 41 сек.	Одиннадцать часов пятнадцать минут сорок одна секунда

Time indications are declined in the nominative/accusative case.

Currencies

Currency indications with $ and DM. are recognized. The currency symbols may precede and follow the numeral.

$40 - 40$	Сорок долларов
25 DM - DM 25	Двадцать пять немецких марок

The currency abbreviations руб., коп. and евр. should follow the digit string.

402 руб. 51 коп.	четыреста два рубля пятьдесят
25 евр.	Двадцать пять евро

The currency signs and abbreviations mentioned above may be used in combination with the following abbreviations for numerals: тыс., млн., млрд. and трлн. These abbreviations should follow the digit string.

$41 тыс.	Сорок одна тысяча долларов
4 млн. DM	четыре миллиона немецких марок
43тыс. евр.	Сорок три тысячи евро
15 млн. руб.	Пятнадцать миллионов рублей

The currency signs and abbreviations can be preceded - and in case of $ and OM also followed - by decimal numbers.

1,4 тыс. руб.	Одна целая четыре десятых тысяч рублей
$2,05 млн.	Две целых пять сотых милллионов долларов

Other currencies are written in full; they follow the digit string.

100 франков	20000 лир	60 фунтов стерлингов

Currency amounts are declined in the nominative/accusative case.

Abbreviations

A standard dictionary includes all common abbreviations. Ambiguous abbreviations are solved based on the linguistic context: the abbreviation г. is pronounced город when followed by alphabetical characters, года when preceded by a date and грамм(-а,-ов) when preceded by a digit string. Similarly, СМ. is pronounced as сантиметр(-а,-ов) when preceded by a string of digits, but смотри in other cases.

г. Москва	Город москва
Вес: 15г.	Вес пятнадцать граммов
Длина: 41 СМ.	Длина: сорок один сантиметр
СМ. Главу 10.	Смотри главу десять

Acronyms

Acronyms are pronounced as normal words.

НАТО	вуз

Initialisms

Initialisms are pronounced letter by letter.

ВМФ	СНГ

Appendix C
BAR CODE
SYMBOLOGIES

This appendix discusses the "symbologies" or bar code standards supported by the IRISPen Executive software and contains some samples of bar codes the IRISPen can read with great ease.

The bar height, the width of the narrow bar, the narrow to wide ratio and the character string the bar code contains are indicated each time. (The narrow to wide ratio does not apply to all types of bar codes.)

Although some general information is given on the various bar code standards, this manual does not intend to be a complete catalog of all there is to know about bar code symbologies.

CODE 39

Code 39 is an alphanumeric, self-checking, variable-length bar code. A character is defined by five black bars and four spaces; three elements are wide, six are narrow. There is a leading and trailing quiet zone, a start and stop character (the asterisk), an optional check character and a gap between the data characters.

This symbology is sometimes called "3 of 9 code", and when a check character is used even "HIBC" for "Health Industry Bar Code". It is primarily used where alphabetic characters are mandatory as you can not only encode the dig-

its, but also the uppercase letters and seven special characters (- . $ / + % and a space).

An extended version of Code 39 allows to encode the full 128 ASCII character set. The "full" ASCII set refers to all characters up to character 128 in the ASCII table set, the "special" and accentuated symbols from 128 onwards such as ç and ó are *not* included. The option "Full ASCII Set" must be enabled with the "Bar Code" button on the toolbar to support this symbology.

TEST1A

Bar height: 0.3", narrow bar width: 12 mils, narrow to wide ratio: 2.5

TEST2A

Bar height: 0.4", narrow bar width: 18 mils, narrow to wide ratio: 2.0

TEST3B

Bar height: 0.5", narrow bar width: 24 mils, narrow to wide ratio: 2.5

The examples below are full ASCII Code 39 bar codes. The option "Full ASCII set" must be enabled to read them correctly.

A_b/1%!

Full ASCII, bar height: 0.3", narrow bar width: 12 mils, narrow to wide ratio: 2.5

Ho*~?

Full ASCII, bar height: 0.4", narrow bar width: 18 mils, narrow to wide ratio: 2.0

Wh #

Full ASCII, bar height: 0.5", narrow bar width: 24 mils, narrow to wide ratio: 2.5

CODE 93

Code 93 is a alphanumeric, self-checking, variable-length bar code. Every character is composed of nine elements arranged in three bars with adjacent spaces.

A character is defined by five black bars and four spaces; three elements are wide, six are narrow. There is a leading and trailing quiet zone, a start and stop character, a termination bar and two check characters. The check characters are verified but filtered from the output.

Code 93 is an extended version of Code 39 - see above - that allows to encode the full 128 ASCII character set. The "full" ASCII set refers to all characters up to character 128 in the ASCII table set, the "special" and accentuated symbols from 128 onwards such as ç and ó are *not* included. Code 93 is also slightly more compact than Code 39.

This standard is used in an alphanumeric context as you can define the complete 128 ASCII character set.

bSe%$&%

Bar height: 0.3", narrow bar width: 12 mils

IRISPen

Bar height: 0.4", narrow bar width: 18 mils

TS0.5/4b

Bar height: 0.5", narrow bar width: 24 mils

INTERLEAVED 2 OF 5

Interleaved 2 of 5 is a numeric, self-checking, variable-length and highly compact bar code. A character is defined by five black bars and five spaces; two black and two white bars are wide, all others bars are white. There is a leading and trailing quiet zone, a start and stop character and an optional check character.

The encoding is "interleaved" as the odd position numbers are encoded in the bars and the numbers in even positions in the spaces. These bar codes must by necessity contain an even number of digits. (The length is often used as a check in itself.)

031225

Bar height: 0.3", narrow bar width: 12 mils, narrow to wide ratio: 2.5

041820

Bar height: 0.4", narrow bar width: 18 mils, narrow to wide ratio: 2.0

1234567890

Bar height: 0.5", narrow bar width: 24 mils, narrow to wide ratio: 2.0

CODABAR

Codabar is a primarily numeric, self-checking and variable-length bar code for up to 16 data characters. There is a leading and trailing quiet zone and a start and stop character (A B C D).

This standard is used primarily for numeric applications; you can encode the digits and six special characters (- . $ / + :).

The digits and the $ and - symbol are printed with one wide bar and one space. The start and stop characters (A, B, C and D) are printed with one wide bar and two white spaces.

A123456B

Bar height: 0.3", narrow bar width: 12 mils, narrow to wide ratio: 2.5

A+56-945B

Bar height: 0.4", narrow bar width: 18 mils, narrow to wide ratio: 2.0

D001100A

Bar height: 0.5", narrow bar width: 24 mils, narrow to wide ratio: 2.5

UPC-A

UPC-A is a fixed-length, numeric bar code that allows to encode 12 characters.

The bar codes are composed of a number system digit (encoded in odd parity), the manufacturer's code (encoded in odd parity), the product code and a check digit. (The user encodes 11 digits.)

The check digit is verified but filtered from the reading result when the option "Verify and Erase Check Digit" is enabled, otherwise, it is read and included in the output. Disable this option to support the variation of UPC which does *not* contain a check digit.

Apart from the check digit, there is a left, center and right guard pattern. The first half of the bar codes uses odd parity digits, the second half uses even parity digits.

UPC stands for "Universal Product Code" - has been designed to uniquely identify a product and its manufacturer.

A version of UPC, UPC-E, is restricted to 6 digits and "zero suppressed" - UPC-E compresses the 12 digit UPC codes into six digits by suppressing the number system digit, trailing zeros in the manufacturer's code and leading zeros in the product code. The IRISPen Executive does not support this variation.

01234567897

Bar height: 0.3", narrow bar width: 12 mils

00112233445

Bar height: 0.4", narrow bar width: 18 mils

01470258036

Bar height: 0.5", narrow bar width: 24 mils

EAN-8

EAN-8 is a numeric, fixed-length bar code for eight characters. The bar codes are composed of two characters for the country code, followed by five data characters and a check digit. (The user encodes seven digits.)

The check digit is verified but filtered from the reading result when the option "Verify and Erase Check Digit" is enabled, otherwise, it is read and included in the output.

Apart from the check digit, there is a left, center and right guard pattern and there are odd and even parity digits.

EAN-8 - EAN stands for "European Article Number" - is used for products where the origin country must be encoded. It is actually a European version of UPC-A.

EAN-13 is a variation of EAN-8 which allows to encode 13 characters.

1234567

Bar height: 0.3", narrow bar width: 12 mils

0321548

Bar height: 0.4", narrow bar width: 18 mils

0354368

Bar height: 0.5", narrow bar width: 24 mils

EAN-13

EAN-13 is a numeric, fixed-length bar code for 13 characters. The bar codes are composed of two characters for the country digit, followed by ten data characters and a check digit. (The user encodes 12 digits.)

The check digit is verified but filtered from the reading result when the option "Verify and Erase Check Digit is enabled, otherwise, it is read and included in the output.

Apart from the check digit, there is a left, center and right guard pattern and there are odd and even parity digits.

EAN-13 - EAN stands for "European Article Number" - is the standard of the publishing industry for the ISBN numbers on books and publications. EAN-13 is actually a European version of UPC-A. See above.

123456789110

Bar height: 0.3", narrow bar width: 12 mils

987654321014

Bar height: 0.4", narrow bar width: 18 mils

95103578920

Bar height: 0.5", narrow bar width: 24 mils

MSI PHARMA

MSI - also called "Plessey" in some literature - is a variable-length, numeric bar code that allows to encode up to 13 characters.

Historically speaking, MSI was one of the first symbologies ever to be developed. Nowadays, MSI is used mainly in the grocery industry for shelf labels and inventory control. This is no surprise, because it is a "robust" bar code that stands up to well to "tear and wear".

Each character of the bar codes consists of 8 elements - 4 bars and 4 spaces. MSI bar codes are not self-checking, the use of a check digit is therefore highly recommended.

The IRISPen Executive supports MSI Pharma, a variation that is used by pharmaceutical companies. Most other bar code readers do not provide support of the MSI symbology.

CODE 128

Support of Code 128, a specialized bar code symbology for the banking world, is not included in the IRISPen Executive software. Contact I.R.I.S. to learn how to obtain this software option.

Appendix D
HANDWRITING AND FORM DESIGN RULES

Where less than optimal reading results are obtained with the handprinting recognition of the IRISPen Executive, it may be useful to adapt your writing style and optimize your forms to maximize your productivity. A few simple rules can teach you why substitutions occur and how to avoid them.

THE WRITING IS IN THE PEN

The first and most important rule is that you should use a sufficiently thick ballpoint. Thin ballpoints do not provide acceptable images: scanning always implies some degrading of the image and zones which aren't sufficiently dark may not be detected. Black pens yield better results than blue pens.

Do not use pencils: the grey images they produce are not always contrasted enough to be picked up by the pen scanner!
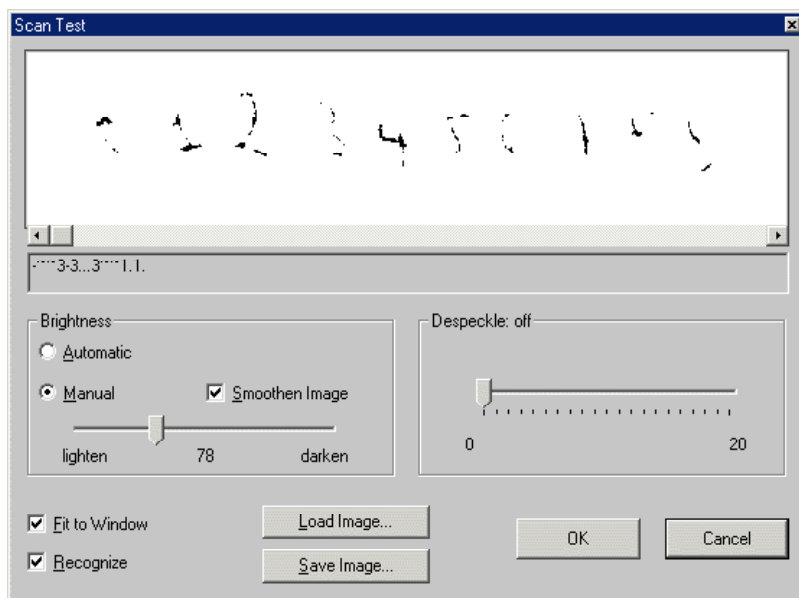
DROP OUT MEANS BACKGROUND

Red and yellow are the IRISPen's invisible or "drop out" colors. This implies that all data in those colors cannot be recognized. On the other hand, these colors can be used efficiently as background color in the blank forms to indicate the various zones to be filled out etc.
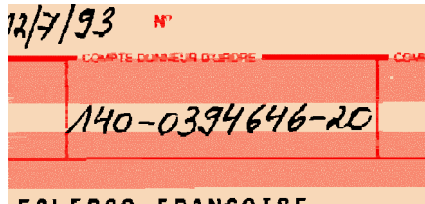
The drop out colors can also be used to ensure spaced printing by having the individual numbers placed in predefined blocks.

DON'T STYLIZE TOO MUCH

Don't add too many loops where none are necessary. Excessively stylized forms may be more difficult to recognize.

The number 2 provides a good example: we evolve from a situation where the form is hardly recognizable to a form where loops are added and make the symbol extremely stylized.

These other digits are also highly stylized: extra loops are added which could lead to rejection or substitution if exaggerated.

LOOPS OPENED AND CLOSED

As extreme stylization can reduce the recognition rate, a simple writing rule could be this: don't open loops which should be closed, don't close loops which should be open!

The samples above show how characters can get confused because loops which should remain open were closed. Taking the examples below, are we dealing with a 4 or a 9, a 4 or a 6 and a 3 or a 5?

HOW TO WRITE ONES

The number 1 can be written in the "American" and "European" style. Ones can be "underlined" or not. This horizontal underlining bar does not have to touch the rest of the font shape.
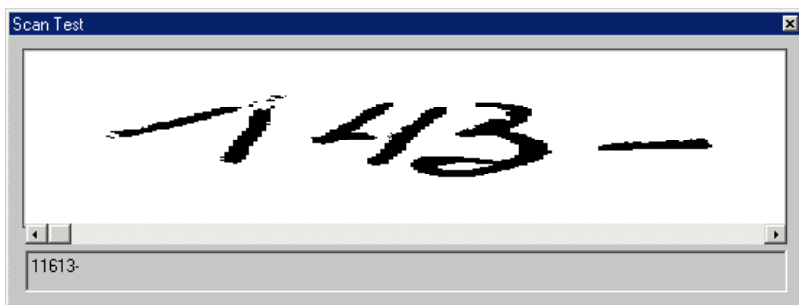
There is no risk of mistaking one for a seven.

ל ר   ר   ד   ך

ל ר   ד   ד   ד

AVOID BROKEN CHARACTERS

Ensure that the symbols 1 and 4 symbol aren't broken when contiguous spacing is selected. For the symbol 1, there is a real risk of recognizing two symbols 1, for 4, you might recognize 61.

This phenomenon only occurs when contiguous spacing is selected, *not* with spaced segmentation.

For the symbol 5 on the other hand, it doesn't matter whether the horizontal bar is separated from the rest of the font form.
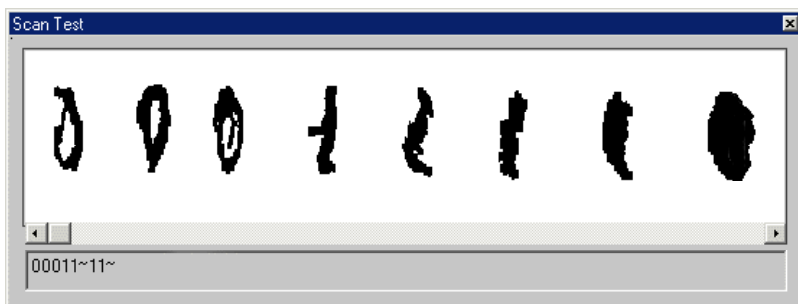
5 5 5 5 5 5

AVOID RETRACING

Avoid retracing as this reduces the image quality and clarity of the handprinted symbols. Symbols which are entirely stricken out are rejected.

In our example, the recognition system makes some lucky guesses about characters which were retraced to correct them, and can only assume that the thin stricken out characters are ones. The ink spot on the extreme right is simply rejected.